Graph-based Representation of Syntactic Structures of Natural Languages based on Dependency Relations

Chang-ho An[1], Zhanfang Zhao[2], Hee-kyungMoon[*3]

[1]Financial Information Engineering, Seokyeong Univ., Seoul, Republic of Korea
[2]Department of Computer Engineering, Hebei GEO Univ., China
[*3]Center for Educational Innovation, Wonkwang Univ., Iksan, Republic of Korea
choan@skuniv.ac.kr[1], zzfsjz@gmail.com[2],ybnjcw@wku.ac.kr[3]

Abstract:Deep Learning approach using probability distribution to natural language processing achieves significant accomplishment. However, natural languages have inherent linguistic structures rather than probabilistic distribution. This paper presents a new graph-based representation of syntactic structures called syntactic knowledge graph based on dependency relations. This paper investigates the valency theory and the markedness principle of natural languages to derive an appropriate set of dependency relations for the syntactic knowledge graph. A new set of dependency relations derived from the markers is proposed. This paper also demonstrates the representation of various linguistic structures to validate the feasibility of syntactic knowledge graphs.

Keywords:Natural Language Processing (NLP), dependency relations, markedness, valency, knowledge graph.

1. Introduction

Linguistic intelligence is one of the ultimate goals of Natural Language Processing (NLP) in Artificial Intelligence (AI). For several decades, a considerable amount of research on language modeling and syntactic/semantic analysis has been accomplished to understand the written texts and spoken dialogs. Recently, a revolutionary approach prompted by Deep Learning (DL) provides a breakthrough insight in NLP and achieves the significant advancement [1, 2, 3, 4, 5]. Several innovative language models using Attention mechanism and Transformer such as ELMo, BERT, and OpenAI GPT demonstrate remarkable performance in text generation, sentiment analysis, question answering, the conversational chatbots, machine translation, and many other important applications of NLP. Significantly, the most recent language model GPT-3 shows human-like, impressive capability in natural language performance [3] However, despite such the noteworthy progress initiated from DL, the substantive issues inherent in natural languages still remain in efficient processing and understanding natural languages. The current language models of NLP are based on a probability distribution over sequences of words [4, 5]. However, what language comprehension is really about is not a probabilistic prediction, but conceptual interpretation. Natural languages are a formal production system holding native syntactic/semantic structures, unlike random probabilistic structures. This fact implicates that language models should stand on natural languages' linguistic perspectives rather than random stochastic events.

In general, computational linguistics has tried to develop computational modeling of natural language, as well as the appropriate computational interpretation of natural language phenomena. For several decades, a considerable amount of research on grammar formalisms has been accomplished from diverse language modeling perspectives. Most approaches have focused on the computational interpretation using diverse mechanisms such as grammar rules, feature-based unification, and logic inference [6, 7]. Although the computational interpretation contributes to understanding natural languages' linguistic properties, it does not provide the expected achievement in linguistic performance.

Nowadays, dependency relations become a common framework in natural language analysis. Since the dependency relations are evident and efficient for syntactic/semantic analysis, this approach can improve NLP applications' linguistic performance. However, there are many variations in dependencies, and they lack a shared consensus on the set of dependency relations. Most of all, there is no definite approach to define dependency relations. In addition to these, NLP approaches using dependency relations use a rigid dependency tree diagram to represent syntactic/semantic structures. The dependency tree is less efficient in representing linguistic knowledge. As a knowledge graph (KG) is used as a general model to represent domain knowledge, the dependency graph is desirable to represent syntactic/semantic structures of natural languages [8, 9]. So, a formal way to define dependency relations based on the universality of natural languages and graph-based representation of linguistic structures is a significant issue in NLP.

This paper presents a graph-based representation of syntactic/semantic structures of natural languages similar to KG. The proposed syntactic knowledge graph is based on dependency relations. This paper investigates the universal principles of natural languages to derive an appropriate set of dependency relations for the syntactic knowledge graph. This paper proposed a new set of dependency relations based on the valency theory and the markedness principle.

This paper also demonstrates the usability of the graph-based representation of syntactic/semantic structures of natural languages.

The remainder of this paper is structured as follows. Section 2 reviews the related work. Section 3 discusses the valency theory and the markedness principle that are the theoretical foundation to derive dependency relations from natural languages' diachronic perspectives. Section 4 analyses the linguistic properties of dependency relations deduced from Section 3. Section 5 presents the construction of syntactic knowledge graphs using the derived dependency relations and markers. Section 5 demonstrates the representation of various linguistic structures to validate the feasibility of syntactic knowledge graphs. Section 6 summarizes the contributions and puts forth the prospects for further work.

2.   Related Work

The main objectives of computational linguistics are to explore syntactic/semantic structures inherent in natural languages. For several decades, a considerable amount of research on grammar formalisms has been accomplished from the diverse perspectives of language modeling. Several notable grammar formalisms such as Lexicalized Functional Grammar(LFG), Categorical Grammar(CG), General Phrase Structure Grammar (GPSG), and Head-driven Phrase Structure Grammar (HPSG) have been developed to describe the complex linguistic structures [6, 7]. These grammar formalisms employ a rule-based system, logic-based system, or feature-based system using unification as the underlying mechanisms. In general, the gramma formalisms adopt syntactic tree structure based on the compositional phrase structures. However, unfortunately, such approaches do not show the expected linguistic analysis achievement, although these grammar formalisms provide linguistic insights into natural languages.

Recently, a revolutionary approach motivated by DL provides a breakthrough insight in NLP and achieves the significant advancement. Several language models such as ELMo, BERT, and GPT-3 shows the remarkable performance in natural language processing [1, 2, 3, 4, 5]. The language models developed in DL demonstrate human-level language performance in a conversational chatbot, sentiment analysis, machine translation, text summarization and generation, and question answering that are deeper applications of NLP. The language prediction approach based on DL generally uses vector semantics based on a probability distribution over sequences of words [4, 5]. Although the NLP approach using distributed, probabilistic semantics demonstrates surprising performance and shows great promise, it still exhibits some issues and arguments that have long plagued DL [10]. Since the language model implemented in this approach is mostly a black box to humans, it does not provide any substantial properties about natural languages. There are no ways for a deeper understanding of how language models work. In other words, there are no formal representation methods to understand syntactic/semantic structures, only vectorised values.

Natural languages are a generative system based on the unique linguistic principles that systematically compose diverse, complex structures. NLP should be able to stand on the linguistic properties of natural languages to exploit and describe linguistic structures. The dependency relations attract considerable attention for syntactic and semantic analyzis in NLP [12]. Nowadays, it is common to use dependency relations in natural language analysis since the dependency relations provide the underlying foundation for representing the linguistic structures [11, 12, 13]. Many systems and open tools, such as Stanford CoreNLP, are widely available to provide the standard framework for developing natural language applications [11, 13].

3.   Valency Theory and Markedness Principle

Natural languages are a kind of generative systems. The linguistics uses grammar rules to describe the generative capability of natural languages. However, the grammar rules to generate linguistic structures are a formal system that enables more fundamental valency property of natural languages. The valency value of linguistic elements plays the principal role in constructing complex linguistic structures. In the realization of linguistic structures, the valency property is closely related to the markedness principle. This section describes the valency theory and linguistic relationship with markedness.

3.1 Valency Theory for Linguistic Structures

The valency theory derived from chemistry regards as the universality of natural language that can clarify the underlying principle of how a sentence is constructed or generated. The origins of valency theory are found in dependency grammar formalism, especially in Lucien Tesnière [14, 15]. Valency theory takes an approach towards linguistic constructions that focus on verbs' syntactic and semantic valencies and occasionally of arguments. In the valency framework, the verb is considered the most central element of a sentence and the major determinant of its structure. The valency is the verb's ability to open up certain positions in its syntactic environment, which can be filled by obligatory or optional complements. The arguments that a verb can take are defined in terms of its valency value. The valency pattern consisting of various types of valency values is a model of a sentence containing a fundamental element (typically, the verb) and a number of dependent elements referred to as arguments, expressions,

or complements whose number and type are determined by the valency pattern of the verb [16, 22]. The following description of (1) is a typical example of the valency pattern, where SCU is a subject complement unit [17, 23].

   (1) valency pattern:
      [SCU]+ call +[NP]+[NP]
     participant patterns:
      AGENT + verb + BENEFICIARY + AFFECTED
      AGENT + verb + AFFECTED + PREDICATIVE

According to the valency pattern of (1), the sentence, "He called her a taxi. He called her a genius" can be composed.

A lot of the research has published the list of valency patterns [16, 17]. Each list of the valency pattern defines its unique complement types, such as INF, WH-CL and V-ing, and semantic roles such as AGENT, LOCATION, and SOURCE [16]. Most of the valency patterns focus on the depiction of dependency relations accompanied between the verb and arguments. However, this approach to define valency patterns neglects the original objectives of the valency concepts. The valency is the perspective of the generation, while dependency relation is for analyzing linguistic structures. The more critical problem is that there are no investigations on how the valency is realized into dependency relation in the surface sentence. This paper addresses the objectification of dependency relations established by the valency in surface sentences.

3.2 Markedness Principle of Linguistic Functions

The valency is the universal linguistic property to combine with other elements in forming phrases and sentences. The valency properties of verbs are closely related to the overall structures of a clause or sentence, in other words, the sentence complements are dependents of the main verb of a sentence or clause [18]. The valency patterns that are directed binary relations are materialized as dependency relations between the governor and the dependent in surface sentences. However, natural language systems need linguistic apparatus to manifest dependencies in the surface structures definitely. The markedness principle, another important universality of natural languages, is used to specify the syntactic/semantic function of the constituents of a surface structure.

Although there are many linguistic perspectives about the markedness, the dependency relations are embodied in markedness. The markedness plays a role in the specification for the constituents of a sentence's syntactic/semantic roles. Specifically, the valency values and dependency relations are the cohesive principles to generate linguistic structures, and markedness is the apparatus to realize the grammatical functions of dependency relations in sentences. While the conventional concepts of markedness have focused on describing the distinctive features of linguistic elements, markedness should be better understood as the bearer of dependency relations. In other words, dependency relations based on valency patterns can be realized by means of the markedness.

   (2) a. He gave his mother's ring to the bride at the wedding.
      b. Einstein assumed that light travels at a constant speed to derive the relativity.

The simple sentences in (2) show that the markedness plays a vital role in representing syntactic/semantic dependencies. Every constituent should have its marker that represents dependency relation and syntactic/semantic linguistic functions. In a broad sense, two types of the markedness can be recognized in the three linguistic levels of words, phrases, and clauses: the explicit markers that imply syntactic/semantic functions such as shown in prepositional phrases and the implicit markers related to the subcategorization of the predicate.

The explicit markers are used to be a syntactic flag that indicates additional linguistic functions. For example, the prepositions TO and AT of (2-a) and THAT and TO-inf of (2-b) are used to represent syntactic functions and semantic roles. The explicit marker plays the binder's role to connect the dependent constituent to the governor in a surface structure. The explicit markers are the principal element to construct complex sentences by expanding the constituents' primary linguistic functions. And more importantly, it should be noted that the explicit markers become the governor of its dependent since the explicit markers define additional syntactic/semantic function of the dependents. In most research about dependent relations, the explicit markers are regarded as an auxiliary dependent of their associated constituent. However, as shown in (2), the syntactic/ semantic functions are decided by the explicit markers TO and AT, not by bride and wedding. The syntactic/ semantic functions of the complex constituents are also decided by the clausal complement markers such as THAT and TO-inf. The explicit markers as the governors realize the graph-based representation of grammatical structures more consistently and semantically.

In principle, all constituents should accompany their markers that expose linguistic functions in surface structures. In the agglutinative languages such as Korean, the markedness principle is strictly kept and offers free

word-order property. However, some languages like English use unmarked constituents. These languages use the word order as the implicit markers that have linguistic functions. The subject, direct object, and indirect object are the specific positions with the implicit marker. In general, the implicit markers are related to the subcategorization of the predicate [19].

The syntactic/semantic functions of the markers can be categorized into two types: government-dependency and attachment-restriction. The government-dependency markers related to subcategorization are used to construct syntactic structures, while the attachment-restriction markers represent the optional modification relationships to impose additional semantic features. However, it should be noted that semantic features of the markers depend on semantic relationships among the governor, marker, and dependent, although some grammar systems like case grammar try to define semantic features of the markers [20]. It is difficult to disclose semantic features by means of the markers all by itself.

Most of the valency patterns define a lot of the markers to formalize the patterns. However, these approaches do not consider the markers' functions as the enabler to realize valency patterns and dependency relations in surface structures. This paper proposes a compact set of markers, as shown in Table 1, considering the linguistic perspectives of the valency and dependencies. The subcat is the implicit markers of subcategorization positions. The empty is another implicit markers used in the positions where are not for subcategorization. The mark is for the relationship between the marker and its associated constituent. xcomp and ccomp are open clausal complement markers and clausal complement markers, respectively.

Table 1. Categorization of Markers and Dependency Relations

Marker Types		Markers	Dependency Relations
Government-Dependency	explicit markers	xcomp ccomp	subj (subject) iobj (indirect object) dobj (direct object) expl (expletive)
	implicit markers	subcat mark	
Attachment-Restriction	explicit markers	gen prep ccomp xcomp aux pass neg conj	mark (dominant by the marker) bind (modifier binding) link (clause linking) pp (prepositional modifier) amod (adjectival modifier) advmod (adverbial modifier) poss (possession modifier) aux (auxiliary verbal)
	implicit markers	empty	pass (passive auxiliary) neg (negation auxiliary)

4.   Analysis of Dependency Relations by Markedness Principle

The dependency relations play a vital role in analyzing syntactic/semantic structures of natural languages. Many open platforms and tools based on dependency relations such as Stanford CoreNLP are widely available to support the efficient development of NLP applications [11, 12, 13]. However, the definition of dependency relations relies on linguistic intuition without a formalized basis. This section discusses the definition and property of dependency relations based on the markedness.

4.1 Definition of Dependency Relations

Most of the dependency relations used in NLP applications are broadly taken from the linguistic analysis of the contemporary linguistic resources such as textbooks, social media messages, and corpus. This approach tries to extract as many as dependency relations that can cover even idiosyncratic structures. The approach is interested in finding new dependency relations and compacting similar relations. There are no definite criteria or systematic approaches to defining dependency relations, although some basic properties in dependency relations such as uniqueness, non-crossing, and acyclic property can verify the consistency [12].

Since the dependency relations are originated from the valency patterns and realized by the markers in the surface structure, the markers are the starting point of the definition of dependency relations. So the dependency relations can be defined in accordance with syntactic/semantic functions of the markers. This paper proposes a set of dependency relations, as in Table 1. The set of dependency relations in this paper is more compact than the set

of Stanford CoreNLP popular in NLP. This paper focuses on the inherent concepts of the markedness principle of natural languages and reflects the diachronic perspectives of linguistic structures. The mark is the dominant relation between the marker and its associated constituent. The link is clausal dependencies, and the bind is modification relations between governor and dependent. This paper does not consider the detailed semantic functions of link and bind that can excessively breed the diverse semantic dependency relations. The semantic dependencies go beyond the scope of syntactic analysis since the markers' primary functions cannot discriminate the semantic dependencies.

4.2 Properties of Dependency Relations

Since the linguistic functions of the markers can be classified into two types, there are two types of dependency relations accordingly. As shown in Table 1, one is the government-dependency type of dependency relations usually related to the subcategorization of the predicate. This type of dependency relation represents the inherent valency capability of the predicate. The dependency relation of subcategorization such as subj, iobj and dobj is actually the placeholder by means of the relative position to the predicate. So they do not imply any linguistic roles of constituents. The semantic interpretation of these dependencies relies on the contextual meaning of the sentence. The other type is attachment-restriction dependencies used to construct complex syntactic structures or attach additional semantic senses. Even though the proposed dependencies are compact compared to other lists of dependency relations, they are based on the inherent conceptions of the markedness and dependency relations. Thus the proposed set is enough to implement the dependency relations related to the modification structures. However, it should be noted that the proposed set is for the representation of dependency relations between constituents, not for the signature of semantic relationships of the constituent. The semantic interpretation of dependency relations demands another level of NLP as shown in (3). Even though all have dependency relation pp marked by IN, their semantic interpretations are different.

(3) a. He is one of the top chefs in Britain. (place)
b. Life in the 19th century was very different from what it is now. (time)
c. Please pay me in cash. (instrument)

In general, the representation of dependency relations, for example, the dependency diagram of Stanford CoreNLP, uses a unidirectional arrow from the governor to dependent [11, 13]. This representation is inadequate to implement the inherent conceptions of dependency relations. Since there are two types of dependency relations, they should be distinguished in the representation. For the attachment-restriction dependencies, they are optional, not required by the governor. In some senses, these dependencies are autonomous relationships initiated by the modifier. So, it is appropriate to use a directed arrow from the modifier to the core constituent. In other words, the arrow of dependency relations goes from the autonomous constituent to the target constituent.

This paper argues that the markers, whether they are government-dependency or attachment-restriction, play the governor's role to its associated constituent since the explicit markers define additional syntactic/semantic function for the associated. The markers are representative of the associated phrasal structure for representing syntactic/semantic functions.

5.   Syntactic Knowledge Graphs of Natural Languages

The conventional NLP is usually based on dependency tree structures. Although syntactic dependency trees can give linguistic intuition, tree structures have some limitations to support flexible syntactic/semantic processing. The graph-based representation is more efficient for data and knowledge modeling as seen in NoSQL databases and KGs. The syntactic knowledge graph can be incorporated with the domain KGs for knowledge processing. The proposed dependency relations are suitable for syntactic knowledge graphs. This section demonstrates various syntactic knowledge graphs based on dependency relations.

5.1 Simple dependency relations

Fig. 1 is a simple syntactic knowledge graph of (4). The dependency relations are represented with dependency_relation/marker format on the edge.

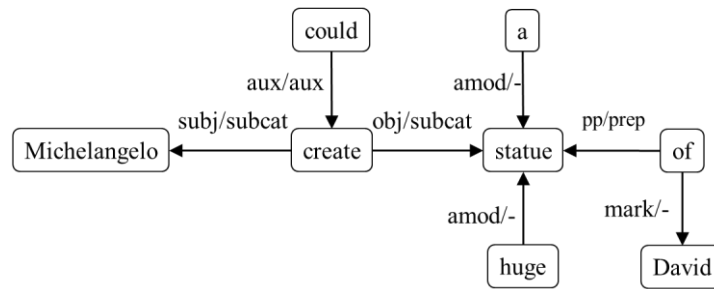(4) Michelangelo could create a huge statue of David.

Fig.1. Syntactic Knowledge Graph with simple dependencies

The syntactic knowledge graphs of Fig. 1 shows that it can localize all linguistic functions. For verbal constituents, theout-going arrows indicate dependent constituents by their valency value. For nominal constituents, the in-going arrows mean syntactic/semantic features. Notably, the syntactic knowledge graph shows the dependencies of each linguistic constituent, not how the constituents are composed.
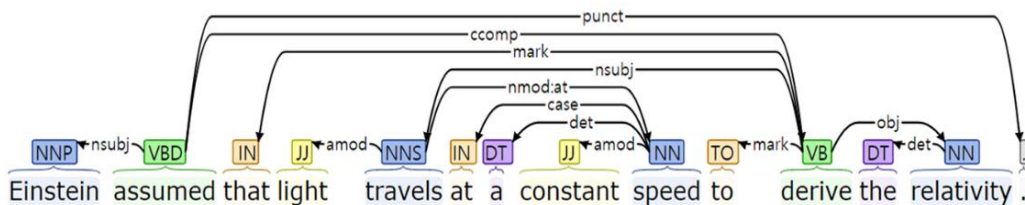
5.2 Dependencies with clausal complements

Fig. 2 is a typical dependency structure with clausal complements, ccomp and xcomp of (5). In the syntactic knowledge graph of (2.a), the predicate assumed dominates the marker THAT with dependency relation dobj/ccomp and the marker THAT governs the predicate travels. This means that the government of the predicate assumed is related to the predicate travels.

(5) Einstein assumed that light travels at a constant speed to derive the relativity.
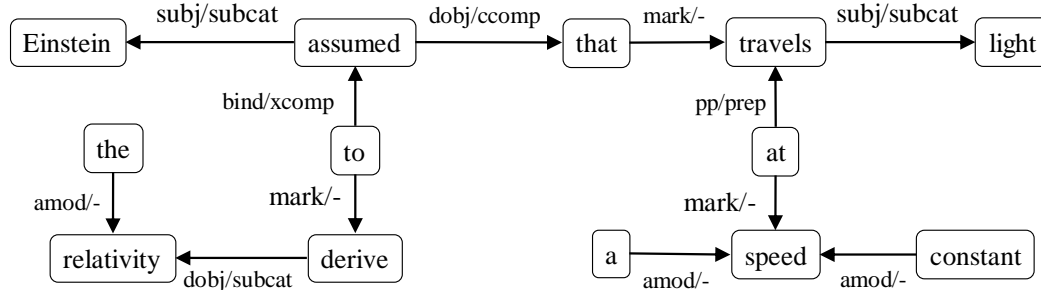
(a) Syntactic knowledge graph with clausal complements

(b) Dependencytree diagram with clausal complements

Fig. 2 Comparison of Syntactic Knowledge Graph and Dependency Tree Diagram

The open clausal complement TO dominating the predicate derive are bound to the predicate assumed. Though the mandatory valency value subj of the predicate derive is unseen, this dependency relationship can be found in the syntactic knowledge graph. Fig. (2.b) is a dependency tree diagram of (5). The dependency relations of the markers TO and THAT dominated by the predicate derive are not explainable. The noun speed dominates different types of constituents, in, a, and constant. The dependency between speed and in is unclear. In general, the dependency tree diagram tries to demonstrate how the constituents are composed in a sentence and to show syntactic structure rather than dependency relations between individual constituents.

5.3 Clausal dependency

Fig. 3 is another example of (6) with an open clausal complement. Two clauses are loosely linked via open complement marker WHEN. The dependency link/ccomp shows that the clause marked by WHEN needs the predicate take, but the predicate take does not.

(6) When I am traveling, I always take something to read in my pocket.
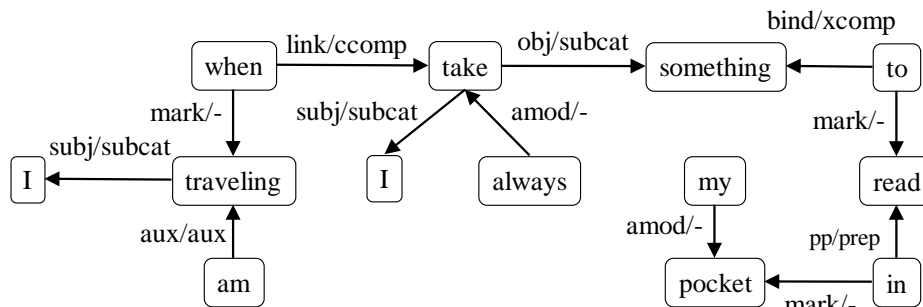

Fig.3. Syntactic Knowledge Graph with Open Clausal Complement

5.4 Long-distance dependency

The linguistic structures involving long-distance dependencies such as topicalization, questions, or relative clauses are the cumbersome problem representing syntactic structure under the current grammar formalisms[21]. Since long-distance dependency occurs when the dependent constituent moves to another place or some constituent intervene dependency relations, it violates the basic syntactic structure rules. Although many resolutions have been proposed, most of them rely on special mechanisms under specific grammar formalisms.

(7) This is the apple that William hit with his arrow.

Fig. 4 is a syntactic knowledge graph of (7) containing a simple relative clause. The predicate hit in (7) has long-distance dependency relation with the antecedent apple. In general, the long-distance dependency is implicitly represented in syntactic representations regardless of their structures. The import issues of long-distance dependency are a reasonable way to restore or estimate unseen dependency relations. In Fig. 4, since the predicate hit does not have mandatory dependency obj/subcat related to subcategorization, the missing dependency can be restored by means of graph traversal. Graph traversal is more efficient than tree traversal or other long-distance dependency algorithms.
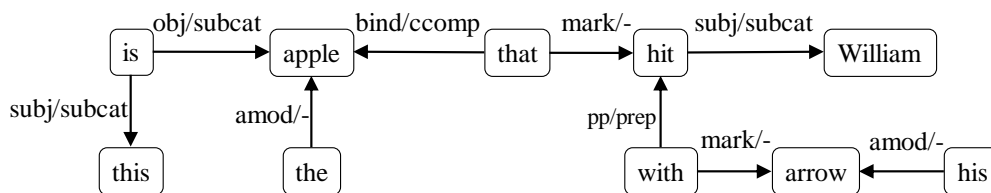

Fig.4. Syntactic Knowledge Graph of Long-Distance Dependency

6.   Conclusions

NLP is the crucial area of AI to realize linguistic intelligence and knowledge processing. Recently, NLP based on Deep Learning achieves breakthrough advancements. However, natural languages have unique linguistics structures inherently, not probabilistic structures. NLP should be able to enlighten linguistic features to achieve more intelligent performance. Nowadays, the dependency relation is the basic framework for natural language analysis and the dependency tree diagram provides essential information for NLP.

However, several issues such as systematic definition and representation of dependency relations remain to be resolved. This paper addresses the syntactic knowledge graph that is a graph-based representation of syntactic/semantic structures based on dependency relations of natural languages similar to knowledge representation and KG. This paper revises the concepts of the valency theory and the markedness principle from

the universality of natural languages. The valency value of the predicate is the underlying capability to generate sentences. In the generation of sentences, the valency pattern is expressed in the form of dependency relations. The dependency relations are embodied in terms of the marker in surface structures. This paper explores the relationships of the valency patterns, dependency relations and markers. The paper then proposes the markers and dependency relations as in Table 1, which is used for syntactic knowledge graphs.

The paper demonstrates a syntactic analysis of various linguistic structures using syntactic knowledge graphs, including clausal complements and long-distance dependencies. This validates that syntactic knowledge graphs are more feasible than dependency tree diagrams in NLP.

7. Acknowledgment

References

1. Ethayarajh, K. (2019). How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings. *arXiv preprint arXiv:1909.00512*.
2. Otter, D. W., Medina, J. R., &Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems.*
3. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Agarwal, S. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165.*
4. Wang, C., Li, M., &Smola, A. J. (2019). Language models with transformers. *arXiv preprint arXiv:1904.09408.*
5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... &Polosukhin, I. (2017). Attention is all you need. *In Advances in neural information processing systems,* 5998-6008.
6. Shieber, S. M. (2003). *An introduction to unification-based approaches to grammar*. Microtome Publishing.
7. Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5), 544-551.
8. Bonatti, P. A., Decker, S., Polleres, A., &Presutti, V. (2019). Knowledge graphs: New directions for knowledge representation on the semantic web (dagstuhl seminar 18371). In *Dagstuhl Reports* (Vol. 8, No. 9). Schloss Dagstuhl-Leibniz-ZentrumfuerInformatik.
9. kumarKaliyar, R. (2015, May). Graph databases: A survey. In *International Conference on Computing, Communication & Automation* (pp. 785-790). IEEE.
10. Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., & Wu, Y. (2016). Exploring the limits of language modeling. *arXiv preprint arXiv:1602.02410.*
11. Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., &McClosky, D. (2014, June). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations* (pp. 55-60).
12. Mel'čuk, I. (2011). Dependency in language-2011. In *Proceedings of International Conference on Dependency Linguistics* (pp. 1-16).
13. De Marneffe, M. C., & Manning, C. D. (2008, August). The Stanford typed dependencies representation. In *Coling 2008: proceedings of the workshop on cross-framework and cross-domain parser evaluation* (pp. 1-8).
14. Nivre, J. (2005). Dependency grammar and dependency parsing. MSI report, 5133(1959), 1-32.
15. Herbst, T., Heath, D., Roe, I. F., &Götz, D. (2013). A valency dictionary of English: *A corpus-based analysis of the complementation patterns of English verbs, nouns and adjective*s (Vol. 40). Walter de Gruyter.
16. Liu, G., & Du, Y. (2019). A Corpus-based Study of Valency Sentence Patterns of English Verbs. *Theory and Practice in Language Studies*, 9(6), 655-665.
17. Faulhaber, S. (2011). *Verb valency patterns: A challenge for semantics-based accounts* (Vol. 71). Walter de Gruyter.
18. Herbert, R. K. (2011). *Language universals, markedness theory, and natural phonetic processes* (Vol. 25). Walter de Gruyter.
19. Roland, D., &Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. *The lexical basis of sentence processing: Formal, computational, and experimental issues*, 4, 325-45.

20. Boas, H. C., & Dux, R. (2017). From the past into the present: From case frames to semantic frames. *Linguistics Vanguard*, 2017, 1-14.
21. Cahill, A., Burke, M., O'Donovan, R., Van Genabith, J., & Way, A. (2004). Long-distance dependency resolution in automatically acquired wide-coverage PCFG-based LFG approximations. *Association for Computational Linguistics*.
22. Atik, G., & Erkan Atik, Z. (2019). Undergraduate Counseling Trainees' Perceptions and Experiences Related to Structured Peer Group Supervision: A Mixed Method Study. Eurasian Journal of Educational Research, 82, 101-120.
23. Bomani, M., Fields, Z., &Derera, E. (2019). The Role of Higher Education Institutions in the Development of SMEs in Zimbabwe. International Journal of Business and Management Studies, 11(2), 1-15.