Language networks and semantic networks

Cite as: AIP Conference Proceedings **2116**, 060007 (2019); https://doi.org/10.1063/1.5114042
Published Online: 24 July 2019

Martin Žáček, and Zdeňka Telnarová

Language Networks and Semantic Networks

Martin Žáček[1, a)], Zdeňka Telnarová[1, b)]

[1] University of Ostrava, Faculty of Science, Department of Informatics and Computers, 30. dubna 22, Ostrava, Czech Republic.

[a)] Corresponding author: Martin.Zacek@osu.cz
[b)] Zdenka.Telnarova@osu.cz

Abstract. How does the human brain work? How can we make sentences and make sentences? These are the questions dealt with already many scientists and several companies developing artificial intelligence. This article presents a study on language networks. At the beginning is the research of the works that have already dealt with this issue. In the next part, the author deals with the application of language networks and semantic networks.

INTRODUCTION

Artificial intelligence modeling is already standard on at least two levels of abstraction, at conceptualization and implementation level. On the first one, i.e. at a higher level of abstraction, conceptualization takes place, the output of which is a conceptual model in language. Although this language has already strict formal rules, it also has language resources that can be understood by the layman as well, thus contributing to the creation of a conceptual model of the defined reference system. The second level uses the language as a rule to implement the model, which is a modification of the first-order logic language and requires deeper mastering of language resources. [1]

In linguistics, the so-called Ullmann's triangle, which illustrates the relationships between reality, abstract concepts and their linguistic representation, is well known (Figure 1). [1]

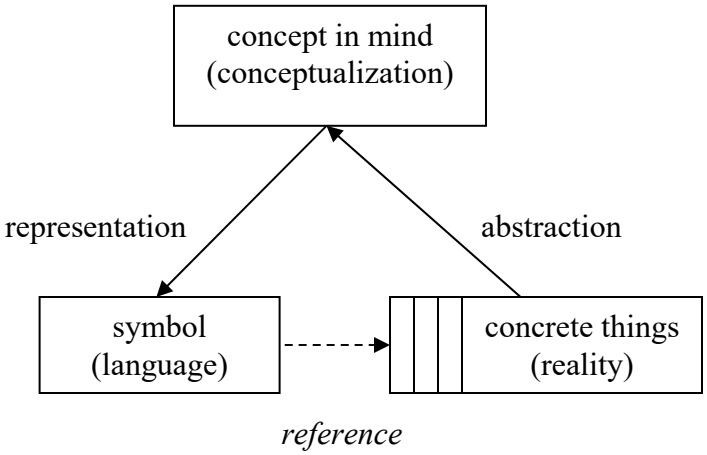

FIGURE 1. The Ullmann's triangle.

International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2018)
AIP Conf. Proc. 2116, 060007-1–060007-4; https://doi.org/10.1063/1.5114042
Published by AIP Publishing. 978-0-7354-1854-7/$30.00

Dashed arrow named references between language and reality shows that a conceptualization always supports the relationship between linguistic expression and the real thing. The language expression of the "college" in the Czech language (in Czech “koleje”) may respond by a conceptualization of reality "railway," in another conceptualization "accommodation buildings for students." [1]

LANGUAGE NETWORKS

If network structure is a potential key for understanding universal statistical trends, then the first step is clearly to define more precisely what kinds of networks are involved. It turns out that there are several possibilities (Fig. 1). [2] First of all, we can look at the network structure of the language elements themselves, and this at different levels: semantics and pragmatics, syntax, morphology, phonetics and phonology. Second, we can look at the language community and the social structures defined by their members. Social networks can help in understanding how fast new conventions propagate or what language variation will be sustained [2, 3].

Moreover, the network organization of individual interactions has been shown to influence the emergence of a self-consistent language [2, 4]

How to build language networks.

Starting from a given text (a) we can define different types of relationships between words, including precedence relations and syntactic relations. In (b) we show them using blue and black arrows, respectively. In figures (c) and (d) the corresponding co-occurrence and syntactic networks are shown. Paths on the network (c) can be understood as the potential universe of sentences that can be constructed with the given lexicon. An example of such a path is the sentence indicated in red. [2]

Nodes have been colored according to the total (in- and out-) word degree, highlighting key nodes during navigation (The higher the degree the lighter its color). In (d) we build the corresponding syntactic network, taking as a descriptive Framework dependency syntax (50), assuming as a criterion that arcs begin in complements and end in the nucleus of the phrase; taking the verb as the nucleus of well-formed sentences. The previous sentence appears now dissected into two different paths converging towards „try.“ An example of the global pattern found in a larger network is shown in (e) which is the occurrence network of a fragment of Moby Dick. In this graph hubs, we have the, a, of, to. [2]

SEMANTIC NETWORKS

Semantic networks (Figure 3) [5, 6, 7] consist of individual words that lexicalize concepts and then map basic semantic relationships (for example such as is special types of isa-relationships) that must be a binary relation (this is the main condition semantic networks). They can be potentially built automatically from corpus data [8, 9, 10, 11]. The topology of these networks reveals highly efficient organizations, where mushrooms are polysemantic words that have a great effect on the overall structure. It has been suggested that hubs organize a semantic web in a categorical representation and explain the ubiquity of polysemy in different languages [8]. In this context, while claiming that polysemy may be some type of historical accident (which should be avoided), the analysis of these sites instead suggests that they are a necessary component of all languages. Additionally, as discussed in [9], the topology without semantic network radiation creates some limitations on how these websites (and the ones previously) can be implemented into neural hardware. The high clustering found in these sites favors search by association, whereas short paths separating two arbitrary items glance [2, 10]

Semantic networks can define different ways. Figure 2 shows a simple network of semantic relationships between lexicalized concepts. Nodes are the terms and links of semantic relations are between concepts. Links are colored to highlight the different nature of relationships. Yellow arcs define isa-relationships (Flower → Rose indicates that Flower is hypernym Rose). Two concepts are linked by a blue arc when there is a partial relationship between them. Binary opposition relations are bi-directional and violet-colored. Isa-relationship defines a tree-structured network, and other relationships create shortcuts that lead the network to display a small world pattern, making it easier and easier to navigate the network. [12, 13]

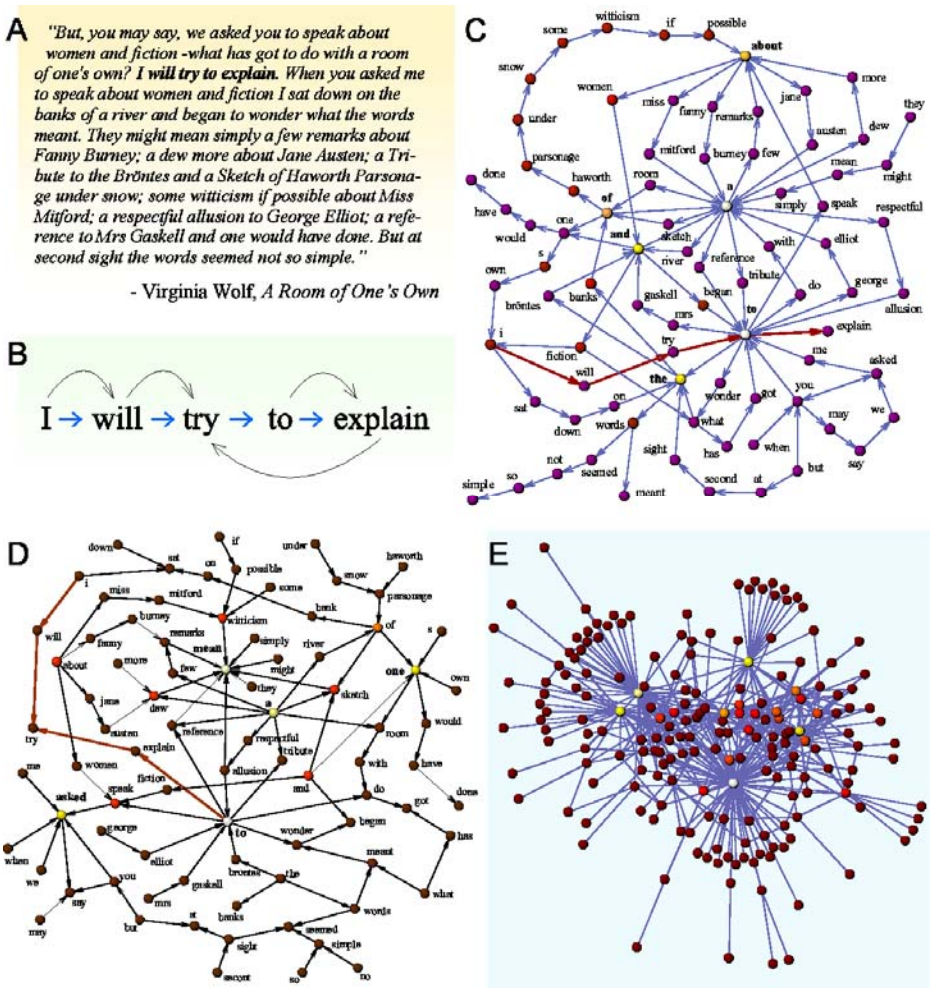

FIGURE 2. The language networks. [2]

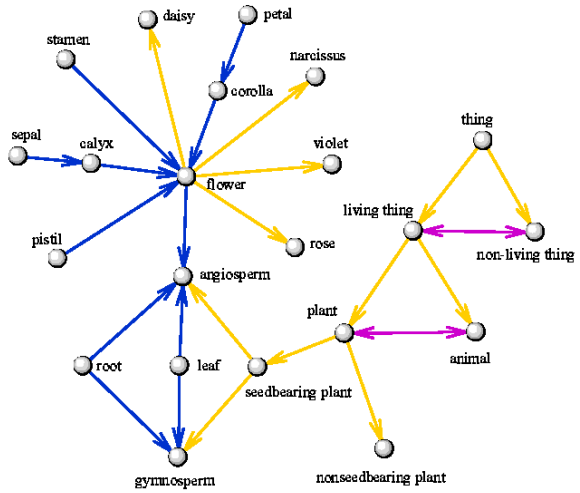

FIGURE 3. The semantic networks. [2]

ACKNOWLEDGMENTS

The research described here has been financially supported by University of Ostrava grant SGS06/PŘF/18. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the sponsors.

CONCLUSION

The aim of the article is to find the meaning of using language networks in artificial intelligence. As we can see in theory, language networks are mainly used in biology and medicine.

This article introduces and compares the language network with semantic networks. In conclusion, we would especially like to thank the authors of Solé, Corominas-Murtra, Valverde, and Steels, who have created a beautiful article: Language structures: Their structure, function, and evolution. This article has become a great inspiration, and for this article I draw the theory from this article.

The authors will use this language network in the Czech language analysis. The Czech language does not have clear rules compared to the English language, so we can create arbitrary sentences without fixed rules.

REFERENCES

1. A. Lukasová, M. Žáček, M. Vajgl, M. and Z. Telnarová. *Formal logic and semantic web* (in Czech), (Plzeň: Polypress s.r.o., 2015). p. 262. ISBN 978-80-261-0408-7.
2. R.V. Solé, B. Corominas-Murtra, S. Valverde and L. Steels, *Language networks: Their structure, function, and evolution.* (Complexity, 2010), 15, pp. 20-26.
3. W. Labov, *Principles of Linguistic change*. Volume II: Social Factors. (Blackwell, Oxford. 2000).
4. B. Corominas and R. V. Solé. *Network topology and self-consistency in language games*. In Journal of Theoretical Biology. 2006, 241(2), pp. 438-441.
5. M. Žáček, R. Miarka, O. Sýkora, Visualization of semantic data, edited by R. Silhavy et al. (Springer, Cham, 2015), pp. 277-285. doi:10.1007/978-3-319-18476-0_28
6. M. Žáček and D. Homola, Analysis of the English morphology by semantic networks, In: International Conference of Computational Methods in Sciences and Engineering 2017, ICCMSE 2017, AIP Conference Proceedings 1906, edited by Simos T.E. et al. (American Institute of Physics, Melville, NY, 2017) p. 080006. doi: 10.1063/1.5012351.
7. A. Lukasová, M. Vajgl and M. Žáček, Knowledge Represented Using RDF Semantic Network. In The Concept of Semantic Web, In: International Conference of Numerical Analysis and Applied Mathematics 2015, ICNAAM 2015, AIP Conference Proceedings 1738, edited by Simos T.E. et al. (American Institute of Physics, Melville, NY, 2015) p. 120012. doi: 10.1063/1.4951895.
8. M. Sigman and G.A. Cecchi, G.A. Global organization of the Wordnet lexicon. (Natl.Acad. Sci. USA, 2002), 99(3), pp. 1742-1747.
9. M. Steyvers, and J.B. Tenenbaum, The largescale structure of semantic networks: statistical analyses and a model of semantic growth. (Cogn. Sci., 2005), 29(1), pp. 41-78.
10. A. Motter, A. et al. Topology of the conceptual network of language. (Phys. Rev., 2002), E65, p. 065102.
11. A.J. Holanda, I. Torres Pisa, O. Kinouchi, A. Souto Martinez, and E. Seron Ruiz, Thesaurus as a complex network. (Physica, 2004), A 344, pp. 530-536.
12. G. A. Miller, The Science of words. Sci. Am. Library. (Freeman and Co. NewYork, 1996).