Review of: "Semblions of Words — The Language of Natural and Artificial Neural Networks"

Olga Chernavskaya[1]

1 P.N. Lebedev Physical Institute of the Russian Academy of Sciences

Potential competing interests: No potential competing interests to declare.

The article gives an ambivalent impression. On the one hand, the questions that the author raises and discusses ¾ what biophysical processes correspond to the functions of consciousness, “Brain & Mind” problem, how memories and associations are realized, etc. – are indeed important, intriguing, and widely discussed problems of modern cognitive science. The work is certainly interesting (“catchy”).

On the other hand, it remains unclear what exactly the author wants to say in this regard. The concept of “Semblion” seems truly new and original, but it is not entirely clear what it is (more on that later).

The second point that causes bewilderment is the following. Despite the declared necessity of a multidisciplinary approach to all the problems discussed, the author uses ONLY natural language: the 10-page article does not contain any formulas, diagrams, or any other attempts to explain the author’s thought using the “metaphors” taken from another area (another language). In other words, it is usually customary to explain one’s own statement (made using natural language, i.e., *words*) by translating it into some other (mathematical or graphic) language. I realize that any object (including an equation, diagram, or model) can be described in words, but WHY?

Further, the author provides a lot of arguments that in the currently popular LLM and ChatGPT models, consciousness and awareness cannot exist, but does not comment in any way on the “*emergence*” effect, i.e., the appearance of something not included in the model initially, which arises for super powerful GPT3, GPT4 as a result of the transition from quantity to quality. But this effect is really interesting and could be related to some (let’s say, artificial) consciousness.

Further, the author constantly refers to the MEM model that he himself had proposed. But what is the purpose of this particular article? One gets the impression (partly confirmed by the author’s comments in the Conclusion) that the article represents a discussion around this model, but 10 pages of verbal arguments in favor of the model are hardly reasonable.

Finally, what is the “Semblion”?

Within the connectionist paradigm, any modern cognitive architecture represents a hierarchical structure of neural layers that is trained in the process of functioning, with the principles of organization may be different. For example, in the NCCA model that we are developing (see, e.g., the paper “To the problem of Digital Immortality”, DOI 10.2139/ssrn.4593718 and refs. therein), we use the so-called “*scaling*” principle. Let me clarify it in more detail.

In NCCA, at the first level of the hierarchy, any external object/situation/phenomenon presented to the system (in the

terminology of the author, this is seemingly called an “external stimulus”) generates a chain of active neurons, with their connections being trained, thereby recording the information (memory) about this object; in our terminology, this chain is called “***image***.” At the next level, due to the specific procedure of competitive interactions, this chain (image) acquires its “***symbol***” (1 neuron). Note in parentheses that at subsequent stages (the next hierarchy layers), this symbol can be associated with a *word*. Semantic information is perceived at intermediate levels, already via the symbols-words.

Further, these symbols themselves can be connected into a certain chain (due to self-organization) and form an “image-of-symbols.” This process is repeated at all hierarchical levels, realizing the *scaling* principle. At the upper levels, complex “generalized images” (*concepts)* appear, with their symbols-concepts being related to *“abstract”* information (not directly connected to any real external stimulus).

In this way, complex multi-level structures of connected neurons are formed. Naturally, the activation of any neuron at the upper layer generates excitation of all those associated with it, which corresponds to memories (and\or imagination), associations, etc. Naturally, the activation of any lower-level neuron generates a reconstruction of the entire corresponding representation, and all those associated with it. Thus, the entire system is “built” from the bottom up and functions by transmitting activity in all directions.

What could be put in correspondence with “Semblion” in this terminology?

I see 2 possible answers.

1) The author considers natural language as a kind of hierarchical structure, in which each*word* represents an analogue to an artificial neuron in a deep learning network. Then, “Semblion” corresponds to any stable formation/combination with sufficiently strong connections, the analogue to our “generalized symbol,” or “image-of-symbols.” For example, the words that make up a well-known poem/phrase form stable strong connections so that at a higher level of the hierarchy, 1 neuron (“Semblion”) appears, representing the symbol of this pattern. Activation of one of these “neurons” (*phonemes*?) provokes activation of the others, i.e., the memories.

2) The author considers a certain cognitive architecture that corresponds to the hierarchically organized layers of artificial neurons, with *some of them* could be treated as ***words***. Then, Semblion corresponds to “symbols-concept” (together with its connections with symbols at lower levels) in our terminology? The mechanism that generates memories is the same as the one described by the author.

Which of these options does the author have in mind? Or is the correct answer some third option? Then this requires some explanation.

In any case, it would be desirable to clarify what exactly the author understands by “Semblion,” using various languages, not only the terms of cognitive linguistics.