Towards Understanding of Natural Language: Neurocognitive Inspirations

Włodzisław Duch[1], Paweł Matykiewicz[1,2], and John Pestian[2]

[1] Dept. of Informatics, Nicolaus Copernicus University, Grudziądzka 5, Toruń, Poland
[2] Department of Biomedical Informatics, Children's Hospital Research Foundation, Cincinnati, OH, USA
{Pawel.Matykiewicz,John.Pestian}@cchmc.org

Abstract. Neurocognitive processes responsible for representation of meaning and understanding of words are investigated. First a review of current knowledge about word representation, recent experiments linking it to associative memory and to right hemisphere synchronous activity is presented. Various conjectures on how meaning arises and how reasoning and problem solving is done are presented. These inspirations are used to make systematic approximation to spreading activation in semantic memory networks. Using hierarchical ontologies representations of short texts are enhanced and it is shown that high-dimensional vector models may be treated as a snapshot approximation of the neural activity. Clustering short medical texts into different categories is greatly enhanced by this process, thus facilitating understanding of the text.

1 Introduction

Low-level cognitive functions involving perception and motor control have reasonable neural models at different level of complexity, from sophisticated spiking neuron biophysical models to quite approximate Hopfield-like and self-organized networks that provide qualitative ideas rather than detailed explanations. Unfortunately, despite great progress in neuroscience, the higher cognitive functions: language, thinking, reasoning, planning, problem solving, creativity, understanding of visual scenes are all poorly understood and lack good working models. Great progress in neuroimaging has not elucidated the precise mechanisms of high-level cognitive functions, because they depend on synchronization of processes at a single neuron or a microcircuit level. Attempts to elucidate such processes at present must be speculative. Even if they prove ultimately too simplistic they may still be fruitful by helping to formulate neurocognitive models of various higher cognitive functions.

In this paper neurolinguistic insights are used to elucidate the process of text understanding and to find useful approximations to the spreading of brain activity during text comprehension. The connectionist approach to natural language has been introduced already in [1], where it was used to explain qualitatively a few linguistic phenomena. The only known system that can deal with linguistic structures is the human brain. The neurocognitive approach to linguistics "is an attempt to understand the linguistic system of the human brain, the system that makes it possible for us to speak and write, to understand speech and writing, to think using language …" [2].

J. Marques de Sá et al. (Eds.): ICANN, Part II, LNCS 4669, pp. 953–962, 2007.

Although this approach has been quite fruitful for understanding neuropsychological, language-related problems, it is relatively unknown in the natural language processing (NLP) community; no practical algorithms for large-scale text analysis have been derived from it.

The basic assumption of neurocognitive computing is that words activate microfeature-based associative networks and that the activation spreads to other parts of the network, which increases the probability of priming dynamic activations of states that facilitate semantic interpretation of words, concepts, sentences and episodes. Basic words and concepts label the action-perception subnetworks, acquiring the meaning directly through references to actions in the environment [3-4]. Constrained spreading activation techniques have recently been applied in information retrieval [5], semantic search techniques [6] and word sense disambiguation [7], although their application is still quite limited.

A brief introduction to the putative neurocognitive processes behind higher cognitive functions is presented in the next section. The section focuses on the use of words and symbols, analysis of priming experiments with pairwise word associations, and recent observations of insight states in the brain. Various approximations of the spreading activation processes in brain networks are discussed and related to the methods used in natural language processing. The challenge is to create approximations that could be used in large-scale, practical NLP projects. An example of how hierarchical ontologies can enhance the representation of short medical texts (summary discharges) illustrates the usefulness of simple approximations. Discussion of the results and their wider implications closes this paper.

2 Representation of Words and Meanings

Linguists have employed symbol manipulation, grammars and parsing techniques, trying to understand languages in conceptual terms. Progress in understanding languages in this way has been rather slow, which has led to the use of statistical techniques to study patterns of language use in large corpora [8]. Although language is based on symbols, logical linguistic analysis may provide only an awkward approximation of the spreading activation and associative processes in the brain. The neurocognitive approach to language draws its inspiration from brain research in trying to understand the processes that make language understanding and production possible.

Sensory systems transform incoming stimuli by extracting from auditory and visual streams such basic quantized elements as phonemes in speech or edges with high contrast in vision. These elementary building blocks form larger patterns, building discrete representations of words and shapes, and in a hierarchical way filling the working memory with information about whole scenes and complex objects, some of them abstract and not even directly related to activation of sensory cortices [9]. The cortex has a layered, modular structure, with columns of about 10^5 densely interconnected neurons, which communicate with other cortical columns in the neighborhood and sometimes also in quite distant areas across the brain, including the opposite hemisphere. Each column contains thousands of microcircuits with different properties (due to the different type of neurons, neurotransmitters and neuromodulators), acting as local resonators that may respond to sensory signals, converting them into intricate patterns of excitations.

Hearing words activates a strongly linked subnetwork of microcircuits that bind articulatory and acoustic representations of a spoken word. Such patterns of activation are localized in most brains in the left temporal cortex, with different word categories coded in the anterior and posterior parts [8-10]. Psycholinguistic experiments show that acoustic speech input is quickly changed into categorical, phonological representation. A small set of phonemes, quantized building blocks of phonological representations are linked together in an ordered string by a resonant state representing word form, and extended to include other microcircuits defining the semantic concept. From the N200 feature of auditory event-related potentials, it has been conjectured that phonological processing precedes semantic activations by about 90 ms [4]. Words seem to be organized in a lexicon, with similar phonological forms activating adjacent resonant microcircuits. Upon hearing a word, a string of connected resonators is activated, creating representation of a series of phonemes that is categorized as a word. Spoken language has a number of syllables and longer chunks of sounds (morphemes) that are strongly associated with each other. They are easily activated when only part of the word is heard, creating the illusion that the whole word has been heard. Categorical auditory perception enables understanding of speaker-independent speech and has clear advantages in a noisy environment, providing speaker-independent speech representation. Strong associations sometimes lead to activation of wrong representations. For example, when only a part of some personal name is heard, often a more common name is substituted.

Phonological representations of words activate an extended network that binds symbols with related perceptions and actions, grounding the meaning of each word in a perception/action network. Various neuroimaging techniques confirm the existence of semantically extended phonological networks, which lends this model of word representation strong experimental support [3,4,10,11]. Symbols in the brain are thus composed of several representations: their sound patterns, pronunciation (vocal motor programs), and their visual and motor associations. This does not resemble the traditional idea of a representation. Learning new concepts prompts minimal changes (convergence) of neural connections that assure unique dynamical states that have the correct relational properties. Hearing a word activates a string of phonemes, increasing the activity (priming) of all candidate words and non-word combinations. A polysemic word probably has a single phonological representation that differs only in its semantic extensions. This encoding automatically ensures that many similarity relations between words, phonological as well as semantic, may automatically be retrieved. Meanings are stored as activations of associative subnetworks that may be categorized and processed further by other areas of the brain. Context priming selects an extended subnetwork corresponding to a unique word meaning, while competition and inhibition in the winner-takes-all processes leaves only the most active candidate networks. The meanings of concepts listed in thesauri or dictionaries are only approximations, because the actual meaning is always modified by the context. Overlapping patterns of brain activations for subnetworks coding word representations lead to strong transition probabilities between the words, and thus to semantic and phonological associations that easily “come to mind”.

During text comprehension, background knowledge stored in the semantic memory is activated, resulting in brain states that contain unique interpretations. Two approaches to knowledge representation prompted by semantic memory are Collins/Loftus spreading activation model [12], and Collins/Quillian's hierarchical semantic memory model [13]. The first has been used in connectionist models of language [1]; the second is the basis for various ontologies. No large-scale semantic networks capturing commonsense knowledge have been built for practical applications, although considerable theoretical work has been done in this area [14,15]. Collecting knowledge for semantic networks that would approximate associative processes in the brain has proved to be quite difficult, since lexical resources such as Wordnet [16] do not contain structural descriptions of concepts. Statistical approaches to context analysis are insufficient in this area because most common sense knowledge is acquired through embodiment and perception, and is so obvious that it is never written down. Recent attempts to analyze machine-readable sources for the creation of large-scale semantic memories have been examined in [17], and the use of word games and active dialogues to extend and correct such knowledge is promising [18]. Ontologies, on the other hand, though they offer taxonomies of concepts [19] that are useful for experts, do not reflect common sense knowledge and lateral associations.

3 Words and Creative Processes

Understanding of words can be regarded as a simple version of problem solving. Recent experiments using the EEG and functional MRI techniques on the "Aha!" insight experience that accompanies some solutions have contrasted insight with analytical problem solving that does not require insight [20,21]. An increased activity in the right hemisphere anterior superior temporal gyrus (RH-aSTG) has been observed during initial solving efforts and during insights. This area is probably involved in higher-level abstractions that can facilitate indirect associations. About 300 ms before insight, a burst of gamma activity was observed. This has been interpreted as "making connections across distantly related information during comprehension (…) that allow them to see connections that previously eluded them" [21]. Bowden *et al.* [20] performed a series of experiments that confirmed the EEG results using fMRI techniques. It is probable that the initial impasse in problem solving is due to the inability of the processes in the left hemisphere, focused on the precise representation of the problem, to make progress. This deadlock is removed when less-focused right hemisphere projects back relevant activations, allowing new dynamical associations to be formed. An emotional component is needed to increase the plasticity of the brain and remember these associations. The "Aha!" experience may thus result from the activation of larger left hemisphere areas by the right hemisphere, with a gamma burst winning the competition for working memory access and thus reaching consciousness. This process occurs more often when the activation of the left hemisphere decreases (giving up conscious efforts to solve the problem), perhaps leading to a short period of knowing that the solution has been found although it has not yet been formulated in symbolic terms. This last step requires synchronization between states in the left hemisphere, defining the transition from the start to the goal through intermediate states.

Such observations may be used as inspirations for neurocognitive models. The LH network codes phonological and visual representations in the visual word form area (VWFA) in the left unimodal occipitotemporal sulcus area. The adjacent lateral inferotemporal multimodal area (LIMA) reacts to both auditory and visual stimulation, and has cross-modal phonemic and lexical links [22]. Extended representations reach to the sensory, motor and premotor cortices [3-4]. Distal connections between the left and right hemispheres require long projections, and therefore neurons in the right hemisphere may generalize over similar concepts and their relations. Most of these RH activations do not have phonological components; the activations result from diverse associations, temporal dependencies and statistical correlations that create certain expectations. For example, hearing the word "left lung" may activate several RH cortical areas that react to all concepts related to lungs and the left side of the upper part of the body, including the heart; hearing "left nose" or "left head" creates a strange feeling. It is not clear what brain mechanism is behind the signaling of this lack of familiarity, but one can assume that interpretation of text is greatly enhanced by "large receptive fields" in the RH, which can constrain possible interpretations, help in the disambiguation of concepts and provide ample stereotypes and prototypes that generate various expectations.

Distributed activations in the right hemisphere also form configurations that activate larger regions of the left hemisphere. High-activity gamma bursts projected to the LH prime its subnetworks with sufficient strength to allow for synchronization of groups of neurons that create distant associations. In problem solving, this synchronization links the initial description D with partial or final solutions S. Such solutions may initially be difficult to justify, they become clear only when all intermediate states T_k between D and S are transversed. If each step from T_k to T_{k+1} is an easy association, a series of such steps is accepted as an explanation. An RH gamma burst activates emotions, increasing the plasticity of the cortex and facilitating the formation of new associations between initially distal states. The same neural processes should be involved in sentence understanding, problem solving and creative thinking.

According to these ideas, approximation of the spreading activation in the brain during language processing should require at least two networks activating each other. Given the word $w = (w_f,w_s)$ with phonological/visual component w_f and extended semantic representation w_s, and the context *Cont*, the meaning of the word results from spreading activation in the left semantic network *LH* coupled with the right semantic network *RH*, establishing a global state $\Psi(w,Cont)$. This state changes with each new word received in sequence, with quasi-stationary states formed after each sentence is understood. It is quite difficult to decompose the $\Psi(w,Cont)$ state into components, because the semantic representation w_s is strongly modified by the context. The state $\Psi(w,Cont)$ may be regarded as a quasi-stationary wave, with its core component centered on the phonological/visual brain activations w_f and with quite variable extended representation w_s. As a result the same word in a different sentence creates quite different states of activation, and the lexicographical meaning of the word may be only an approximation of an almost continuous process. To relate states $\Psi(w,Cont)$ to lexicographical meanings, one can clusterize all such states using dendrograms and use different cutoffs to define prototypes for different meanings.

4 Approximations to Brain States

The high-dimensional vector model of language is a very crude approximation that does not reflect essential properties of the perception-action-naming activity of the brain [3-4]. The process of understanding words (spoken or read) starts from activation of the phonological or grapheme representations that stimulate networks containing prior knowledge used for disambiguation of meanings. This continuous process may be approximated through a series of snapshots of microcircuit activations $\phi_i(w,Cont)$ that may be treated as basis functions for the expansion of the state $\Psi(w,Cont) = \Sigma_i\, \alpha_i\, \phi_i(w,Cont)$, where the summation extends over all microcircuits that show significant activity resulting from presentation of the word w. The high-dimensional vector model used in NLP measures only the co-occurrence of words $\mathbf{V}_{ij} = \langle \mathbf{V}(w_i),\mathbf{V}(w_j)\rangle$ in some window, averaged over all contexts. A better approximation of the brain processes involved in understanding words should be based on the overlap between waves $\langle\Psi(w_1,Cont) \mid \Psi(w_2,Cont)\rangle = \Sigma_{ij}\, \alpha_i\, \alpha_j\, \langle\phi_i(w_1,Cont) \mid \phi_j(w_2,Cont)\rangle$ that depends on time. Systematic study of transformations between the two bases: activation of microcircuits ϕ_i and activation of complex patterns $\mathbf{V}(w_i)$, has not yet been done. The use of waves to describe states makes this formalism similar to that used in quantum mechanics, although no real quantum effects are implied here.

Spreading activation in semantic networks should provide enhanced representations that involve concepts not found directly in the text. Approximations of this process are of great practical and theoretical interest. The model should reflect activations of various concepts in the brain of an expert reading such texts. A few crude approximations to this process may be defined. First, semantic networks that capture many types of relations among different meanings of words and expressions may provide space on which words are projected and activation spread. Each node w in the semantic network represents the whole state $\Psi(w,Cont)$ with various contexts clusterized, leading to a collection of links that capture the particular meaning of the concept. Usually only the main differences among the meanings of the words with the same phonological representation are represented in semantic networks (meanings listed in thesauruses), but the fine granularity of the meanings resulting from different contexts may be captured in the clusterization process and can be related to the weights of connections in semantic networks. The spreading activation process should involve excitation and inhibition, and "the winner takes most" processes. Current models of semantic networks used in NLP are only vaguely inspired by the associative processes in the brain and do not capture such details [14,15].

Quite crude approximation to the spreading activation processes leads to enhancement of the initial text being analyzed by adding new concepts linked by semantic or hierarchical ontological relations. Inhibition between concepts arising from the same phonological word forms should then lead to formation of graphs of consistent concepts, applied recently to disambiguate concepts in medical domain [23]. The enhanced representations are very useful in document clusterization and categorization, as is illustrated using short medical texts in the next section. Vector models may be related to semantic networks by looking at snapshots of the activation of nodes after several steps of spreading the initial activations through the network. In view of the remarks about the role of the right hemisphere, larger "receptive fields" in the linguistic domain should be defined and used to enhance text representations. This is much more difficult because many of these processes have no phonological component and thus have representations that are less constrained and have no directly identifiable meaning. Internal representations formed by neural networks are also not meaningful to us, as only the final result of information processing or decision making can be interpreted in symbolic terms. Defining prototypes for different categories of texts, clusterizing topics or adding prototypes that capture some *a priori* knowledge useful in document categorization [24], is a process that goes in the same direction.

Relationships between creativity and associative memory processes have been noticed long ago [25]. Further experimental support for the ideas described above may be found in pairwise word association experiments using different priming conditions. In [26] puzzling results from using nonsensical words were observed for people with high compared to those with low creativity levels. Analysis of these experiments provided in [27] reinforces the idea that creativity relies on associative memory, and in particular on the ability to link distant concepts together. Adding neural noise by presenting nonsensical words in priming leads to activation of more brain circuits and facilitates in a stochastic resonance-like way a formation of distal connections for not obvious associations. This is possible only if weak connections through chains involving several synaptic links exist, as is presumably the case in creative brains. For simple associations the opposite effect is expected, with strong local activations requiring longer times for the inhibitory processes to form consistent interpretations. Such experiments show that some effects cannot be captured at the symbolic level. It is thus quite likely that language comprehension and creative processes both require subsymbolic models of neural processes realized in the space of neural activities, reflecting relations in some experiential domain, and therefore cannot be modeled using semantic networks with nodes representing whole concepts. Recent results on creation of novel words [27] give hope that some of this process can be approximated by statistical techniques at the morphological level.

5 Visualization of Semantic Similarity

The time-dependent state of the brain $\Psi(w_i,Cont)$ that arises after reading or hearing texts that are understood by the experts should show high similarity for documents of the same category and should be different if documents from other categories are processed. Documents have usually quite sparse representation; for example, hospital discharge summaries by different specialties, but for the same disease, may use completely different vocabularies. Therefore, agglomerative hierarchical clustering methods will show a poor performance in document clustering. The simplest extension is to replace single words (terms) by associations based on synonyms, for example by using the Wordnet synsets [16]. This simulates some of the spreading activation processes in the brain increasing the similarity of documents that use different words to describe the same topic. However, synsets are not useful for very specific concepts that have no synonyms, such as medical concepts used in discharge summaries. To avoid problems with shared common words, only specific concepts that belong to selected semantic types may be used – the process presumably facilitated by the RH.

A better approximation to spreading activation in brain networks is afforded by soft evaluation of the similarity of different terms. Distributional hypothesis assumes that similarity of terms results from similar linguistic contexts [8]. However, in the medical domain and other specialized areas it may be quite difficult to estimate similarity reliably on the basis of co-occurrence, because there are so many specific concepts that there will never be sufficient data to do that. Statistical approaches cannot replace systematic, structured knowledge describing medical concepts. To illustrate that process, two-steps of spreading activation have been made in a network built from ontological relations found in the Systematized Nomenclature of Medicine –Clinical Terms (SNOMED CT) section of the National Library of Medicine's Unified Medical Language System (UMLS) [19]. Discharge summaries for 10 initial diagnoses are represented by a carefully selected semantic feature space (described in [24]). Figs. 1-3 show Multidimensional Scaling (MDS) visualization of records from three strongly overlapping classes only to improve legibility: pneumonia (class 1, 609 records), juvenile rheumatoid arthritis (class 6, 41 records) and otitis media (class 9, 493 records).

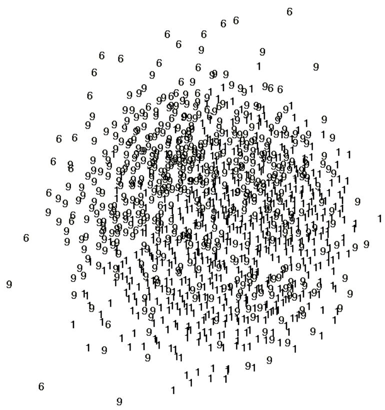

Fig. 1. MDS for original data

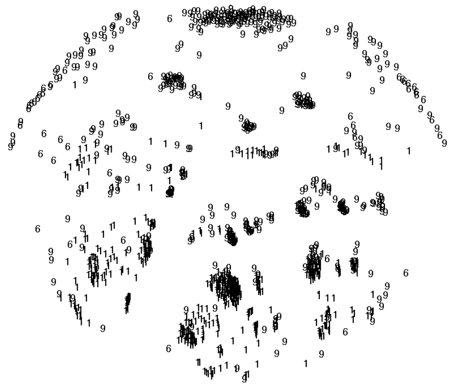

Fig. 2. MDS after first enhancement

Initial feature space is composed from 488 SNOMED CT concepts with high feature-class correlation coefficient (CC>0.5). Visualization of these documents using multi-dimensional scaling (Fig. 1) shows great mixing of documents. A single step of spreading activation through the network, followed by feature selection based on CC > 0.27 extends the feature space to 761 concepts. MDS in this space (Fig. 2) already shows a clear cluster structure. The second iteration with CC>0.5 increases the space to 1138 features and shows even more detailed and fine-grained structure, identifying different subclusters within each category (Fig. 3). For example, bacterial infections may come from Yersinia, Salmonella, Streptococcal and other infections, increasing similarity of all diseases caused by bacteria. In the extended spaces accuracy of classification is also greatly improved – for the 3 classes presented here from about 81% to 87% and 88±4% in crossvalidation tests using linear SVM (for the 10-class case the improvement is on more than 20%). Even quite simple approximations of the spreading of neural activation leads to a significantly improved accuracy in classification.

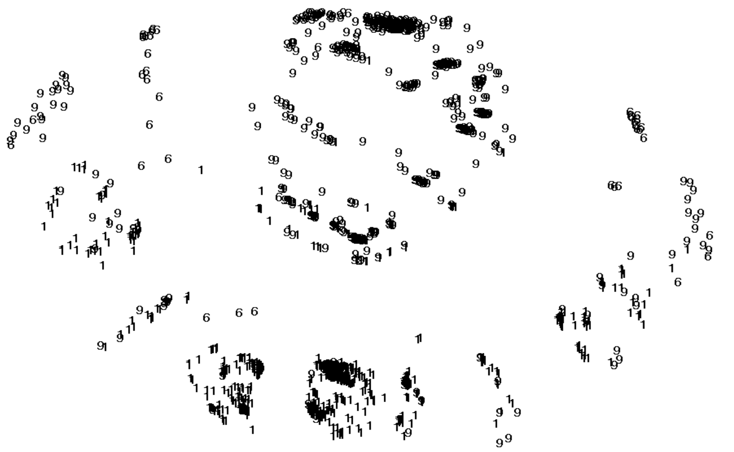

Fig. 3. MDS on medical discharge summaries after two enhancement steps

6 Conclusion

Although linguistic processes are not yet completely understood, following neurolinguistic inspirations may be quite fruitful, allowing one to formulate some crude models of the processes that are responsible for text understanding in real brains. Various approximations to the putative brain processes responsible for language comprehension have been considered, leading to useful algorithms for text analysis. Vector representations of concepts may be regarded as a snapshot of activity patterns, defining connections with other concepts. Relations between spreading activation in neural and in semantic networks, and the vector model of concepts have been elucidated. The role of the right hemisphere, which constrains and guides the spreading activation processes by providing "large receptive fields" for concepts, has been discussed.

It is perhaps surprising that even a crude approximation using two steps of spreading activation with feedback loops leads to such good clusterization and to great improvement in classification on a very difficult problem of summary discharge categorization [24]. Background knowledge has been derived here from synsets, statistical co-occurrences and ontologies. In [24] prototypes of concepts representing *a prio*ri medical knowledge were used, providing crude approximation of the activity of neural cell assemblies in the brain of a medical expert who thinks about a particular disease. Creating numerical representations of various concepts that may be useful in large-scale NLP applications is an interesting challenge. Neurocognitive inspirations lead here to many ideas that will be explored in future work.

Acknowledgment. Support by the Polish Committee for Scientific Research, research grant 2005-2007, and Ohio's Third Frontier is gratefully acknowledged.

References

1. Rumelhart, D.E., McClelland, J.L. (eds.): Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1 & 2. MIT Press, Cambridge (1986)
2. Lamb, S.: Pathways of the Brain: The Neurocognitive Basis of Language. J. Benjamins Publishing, Amsterdam & Philadelphia (1999)

3. Pulvermüller, F.: The Neuroscience of Language. On Brain Circuits of Words and Serial Order. Cambridge University Press, Cambridge, UK (2003)
4. Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R.: Brain signatures of meaning access in action word recognition. Journal of Cognitive Neuroscience 17, 884–892 (2005)
5. Crestani, F.: Application of Spreading Activation Techniques in Information Retrieval. Artifical Intelligence Review 11, 453–482 (1997)
6. Crestani, F., Lee, P.L.: Searching the web by constrained spreading activation. Information Processing & Management 36, 585–605 (2000)
7. Tsatsaronis, G., Vazirgiannis, M., Androutsopoulos, I.: Word Sense Disambiguation with Spreading Activation Networks Generated from Thesauri, 20th Int. Joint Conf. in Artificial Intelligence, Hyderabad, India, pp. 1725–1730 (2007)
8. Manning, C.D, Schütze, H.: Foundations of Statistical Natural Language Processing. MIT Press, Cambridge, MA (1999)
9. Caplan, D., Waters, G.S.: Verbal working memory and sentence comprehension. Behavioral and Brain Sciences 22, 77–94 (1999)
10. Damasio, H., Grabowski, T.J., Tranel, D., Hichwa, R.D., Damasio, A.R.: A neural basis for lexical retrieval. Nature 380, 499–505 (1996)
11. Martin, A., Wiggs, C.L, Ungerleider, L.G, Haxby, J.V: Neural correlates of category-specific knowledge. Nature 379, 649–652 (1996)
12. Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. Psychological Reviews 82, 407–428 (1975)
13. Collins, A.M., Quillian, M.R.: Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior 8, 240–247 (1969)
14. Sowa, J.F. (ed.): Principles of Semantic Networks: Explorations in the Representation of Knowledge. Morgan Kaufmann Publishers, San Mateo, CA (1991)
15. Lehmann, F.(eds.): Semantic Networks in Artificial Intelligence. Pergamon, Oxford (1992)
16. See http://wordnet.princeton.edu
17. Szymanski, J., Sarnatowicz, T., Duch, W.: Towards Avatars with Artificial Minds: Role of Semantic Memory. Journal of Ubiquitous Computing and Intelligence (in press)
18. Szymanski, J., Duch, W. (eds.): Semantic Memory Knowledge Acquisition Through Active Dialogues. Int. Joint Conf. on Neural Networks (in press). IEEE Press, Orlando (2007)
19. UMLS Knowledge Sources, available at http://www.nlm.nih.gov/research/umls
20. Bowden, E.M., Jung-Beeman, M., Fleck, J., Kounios, J.: New approaches to demystifying insight. Trends in Cognitive Science 9, 322–328 (2005)
21. Jung-Beeman, M., Bowden, E.M., Haberman, J., Frymiare, J.L., Arambel-Liu, S., Greenblatt, R., Reber, P.J.: Neural activity when people solve verbal problems with insight. PLoS Biology 2, 500–510 (2004)
22. Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., Cohen, L.: Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. Neuron 50, 19–204 (2006)
23. Matykiewicz, P., Duch, W., Pestian, J.: Nonambiguous Concept Mapping in Medical Domain, Lecture Notes in Artificial Intelligence. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2006. LNCS (LNAI), vol. 4029, pp. 941–950. Springer, Heidelberg (2006)
24. Itert, L., Duch, W., Pestian, J.: Medical document categorization using a priori knowledge, Lecture Notes in Computer Science. In: Duch, W., Kacprzyk, J., Oja, E., Zadrożny, S. (eds.) ICANN 2005. LNCS, vol. 3696, pp. 641–646. Springer, Heidelberg (2005)
25. Mednick, S.A.: The associative basis of the creative process. Psychological Review 69, 220–232 (1962)
26. Gruszka, A., Nęcka, E.: Priming and acceptance of close and remote associations by creative and less creative people. Creativity Research Journal 14, 193–205 (2002)
27. Duch, W., Pilichowski, M.: Experiments with computational creativity. Neural Information Processing – Letters and Reviews 11, in press (2007)