

Experiments with Computational Creativity

Włodzisław Duch^{1,2} and Maciej Pilichowski¹

¹Dept. of Informatics, Nicolaus Copernicus University, Grudziądzka 5, Toruń, Poland,

²School of Computer Engineering, Nanyang Technical University, Singapore

Google: Duch; macias@mat.umk.pl

(Submitted on December 1, 2003)

Abstract—Neurocognitive model inspired by the putative processes in the brain has been applied to invention of novel words. This domain is proposed as the simplest way to understand creativity using experimental and computational means. Three factors are essential for creativity in this domain: knowledge of the statistical language properties, imagination constrained by this knowledge, and filtering of results that selects most interesting novel words. These principles are implemented using a simple correlation-based algorithm for auto-associative memory that learns the statistical properties of language. Results are surprisingly similar to those created by humans. Perspectives on computational models of creativity are discussed.

Keywords—Creativity, brain, language processing, higher cognitive functions, neural modeling.

1. Introduction

Many low-level cognitive functions involving perception and motor control have already reasonable neural models, but higher cognitive functions (language, thinking, reasoning, planning, problem solving, understanding of visual scenes) are still poorly understood. Creativity seems to be one of the most mysterious aspects of the human mind. As with many concepts related to human behavior creativity is not easy to define or measure. Research on creativity has been so far pursued by philosophers, educators and psychologists, who described stages of creative problem solving and devised various tests that can be used to assess creativity. MIT Encyclopedia of Cognitive Sciences [1], Encyclopedia of Creativity [2] and Handbook of Human Creativity [3] do not mention brain mechanisms or computational models of creative processes. Thus any attempt to elucidate brain processes behind creative thinking at present has to be speculative.

Creativity has been defined by Sternberg [3] as “the capacity to create a solution that is both novel and appropriate”. In this sense creativity manifests itself not only in creation of novel theories or inventions, but permeates our everyday actions, understanding of language and interactions among people. This becomes quite clear when interactive systems, such as chatterbots, avatars or robots attempt to simulate human behavior. High intelligence is not sufficient for creativity although it is quite likely that both have similar neurobiological basis. Arguably the simplest domain in which creativity is frequently manifested is in the invention and understanding of novel words. This ability is shown very early by babies learning to speak and understand words. Neurocognitive approach to use of words and symbols should draw inspirations from experimental psychology and brain research and help to understand putative brain processes responsible for creativity in the domain of novel word creation. This could be a good area for more precise tests of creative processes using computational, theoretical and experimental approaches. A first step in this direction is presented in this paper.

Neural networks find applications at the perceptual, low-cognition level, while creativity is regarded as a high-level cognitive process. Artificial intelligence has focused only on symbol manipulation for problem solving and it is rather doubtful that this type of approach can lead to creative behavior. Symbolic language may provide only an awkward approximation to spreading activation and associative processes in the brain. At the level of higher cognitive processes pattern formation in the brain is very rapid, with words and concepts labeling the action-perception subnetworks [4]-[7]. A neurocognitive model of brain processes should link low-level and

higher-level cognitive processes, and allow for analysis of relations between mental objects, showing how neurodynamical processes are manifested in inner experience at the psychological level. A fruitful way to look at this problem [8] is to start with the neurodynamical description of brain processes and look for approximations to the evolution of brain states in low-dimensional space where each dimension may be related to inner experience. This idea has been used to model category learning in experimental psychology, showing why counter-intuitive answers may be given in some situations [9].

In the next section relevant facts that justify our computational model are presented. Algorithms used to generate novel words are introduced in section three, and some results are presented in section four. Discussion of these results and future directions of the neurocognitive approach to creativity closes this paper.

2. Creativity, Words and Brain Processes

The “g-factor” used to measure intelligence is highly correlated with working memory capacity, perceptual speed, choice and discrimination reaction times, the structure of EEG event-related potentials (ERP potentials), nerve conduction velocity, and cerebral glucose metabolic rate during cognitive activity [10]. Brains of creative and intelligent people probably differ in the density of synaptic connections, contributing to the richer structure of associations, and more complex waveforms of the ERP potentials. Although gross brain structures are identical in all normal infants there is a lot of variability at the microscale that may result in exceptional talents or developmental disorders. Auditory event-related potentials (ERPs) for consonant sounds, recorded from left or right hemisphere less than two days after birth, predicted with over 80% accuracy the level of reading performance of children eight years later [11][12]. The structure of these potentials depends on the speed of neural signal transmission and the density of neural connections.

Sensory systems transform the incoming stimuli extracting from auditory and visual streams basic quantized elements, such as phonemes or edges with high contrast. These elementary building blocks form larger patterns, building discrete representations for words and shapes, and at the working memory level of whole scenes and complex abstract objects. Current working memory model of Baddley [13] includes central executive functions in the frontal lobes, focusing, switching and dividing attention, and providing an interface to the long-term associative memory via episodic memory buffer. In this model working memory includes also two slave short-term memory buffers, auditory phonological loop and visuospatial sketchpad. Individual differences in the ability to control attention depend on the working memory capacity [14] and are obviously reflected in the ability to solve problems and exhibit intelligent behavior. Electrophysiological and brain imaging studies show the involvement of frontal, temporal and parietal associative cortex in storage of working memory information. However, it is quite likely that working memory is not a separate subsystem, but simply an active part of the long-term memory (LTM) network (see the review of the evidence for this point of view in [15]) due to priming and spreading of neural activation. The same brain regions are involved in perception, storage and re-activation of LTM representations. Some activated LTM subnetworks may be in the focus of attention of the central executive (presumably this is the part we are conscious of) and some may be activated but outside of this focus.

To understand the higher cognitive processes one should start from understanding how symbols are stored and used in the brain. Cortex has layered modular structure, with columns of about 10^5 densely interconnected neurons, communicating 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 the sensory signals converting them into intricate patterns of excitations. Hearing words activates a strongly linked subnetwork of microcircuits that bind articulatory and acoustic representation of a spoken word. Such patterns of activation are localized in most brains in the left temporal cortex, with different word categories coded in anterior and posterior parts [16]-[18]. The ability to hear and recognize a word does not imply that it can be properly pronounced. Production of words requires precise motor programs that are linked to phonological representations in temporal cortex, but are stored in frontal motor cortex (Broca's area), connected to the temporal areas via a bundle of nerve fibers called the arcuate fasciculus. Damages (lesions) to this fiber or to the cortex processing auditory and language-related information leads to many forms of aphasia [19].

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 is linked together in an ordered string by a resonant state representing word form, and extended to include other microcircuits defining 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. Words seem to be

organized in a lexicon, with similar phonological forms activating adjacent resonant microcircuits. Upon hearing a word 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) strongly associated with each other. They are easily activated when only part of the word is heard, creating illusion that the whole word has been heard properly. Categorical auditory perception enables understanding of a speaker-independent speech and is more reliable in a noisy environment but strong associations sometimes lead to activation of wrong representations. For example, hearing part of the person's name frequently more common name is noted or remembered.

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 existence of semantically extended phonological networks, giving this model of word representation strong experimental support [4]-[7]. Symbols in the brain are thus composed of several representations: how they sound like, how to say them, what visual and motor associations they have. This encoding automatically assures that many similarity relations, phonological as well as semantic, between words 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. Hearing a word activates string of phonemes increasing the activity (priming) of all candidate words and non-word combinations (good computational models of such phenomena in phonetics are described in [21]). Polysemic words probably have a single phonological representation that differs only by their semantic extensions. Context priming selects 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 subtle meaning of a concept, as stored in dictionaries, can only be approximate, as it is always modified by the context. Overlapping patterns of brain activations for subnetworks coding word representations lead to strong transition probabilities between these words and thus semantic and phonological associations that easily "come to mind".

Relationships between creativity and associative memory processes have been noticed a long time ago [22]. The pairwise word association technique is perhaps the most direct way to analyze associations between subnetworks coding different concepts. These associations should differ depending on the type of priming (semantic or phonological), structure of the network coding concepts, the activity arousal due to the priming (the amount of energy pumped into the resonant system). In a series of experiments [23] phonological (distorted spelling) and semantic priming was applied, showing for a brief (200 ms) moment the priming cue (word) before the second word of the pair was displayed. Two groups of people, with high and low scores in creativity tests were participating in this experiment. Two type of associations were presented, close and remote, and two types of priming, positive (either phonological or semantic relation to the second word) and neutral (no relation). Creative people should have greater ability to associate words and should be more susceptible to priming. Less creative people may not be able to make remote associations at all, while creative people should show longer latency times before noticing such associations or claiming their absence. This is indeed observed, but other results have been quite puzzling [23]. Neutral priming, based on the nonsensical or unrelated words, increased the number of claims that words are related, in case of less creative people stronger than positive priming, and in case of more creative people in a slightly lower way. Phonological priming with nonsensical sounds partially activates many words, adding intermediate active configurations that facilitate associations. If associations between close concepts are weak neutral priming may activate intermediate neural oscillators (pumping energy to the system, increasing blood supply), and that should help to establish links between paired words, while positive priming activates only the subnetwork close to the second word, but not the intermediate configurations. For creative people close associations are easy to notice and thus adding neutral or positive primes has similar small effect. Situation is quite different for remote associations. Adding neutral priming is not sufficient to facilitate connections in less creative brains when distal connection are completely absent, therefore neutral priming may only make them more confused. Adding some neural noise may increase the chance to form resonance state if weak connections exist in more creative brains – in the dynamical systems language this is called the stochastic resonance phenomenon [24]. On the other hand adding positive priming based on spelling activates only phonological representations close to that of the second word, therefore there is no influence. Priming on positive (related) meaning leads to much wider activation, facilitating associations. These results support the idea that creativity relies on the associative memory, and in particular on the ability to link together distant concepts.

Creativity is a product of ordinary neurocognitive processes and as such should be amenable to computational modeling. However, the lack of understanding what exactly is involved in creative activity is one of the main reasons for the low interest of the computational intelligence community in creative computing.

Problems that require creativity are difficult to solve because neural circuits representing object features and variables that characterize the problem have only weak connections, and the probability of forming appropriate sequence of cortical activities is very small. The preparatory period – reading and learning about the problem – introduces all relevant information, activating corresponding neural circuits in the language areas of the dominant temporal lobe, and recruiting other circuits in the visual, auditory, somatosensory and motor areas used in extended representations. These brain subnetworks are now “primed”, and being highly active reinforce mutually their activity, forming many transient configurations and inhibiting at the same time other activations. Difficult problems require long incubation periods that may be followed by an impasse and despair period, when inhibitory activity lowers activity of primed circuits, allowing for recruitment of new circuits that may help to solve the problem. In the incubation period distributed sustained activity among primed circuits leads to various transient associations, most of them short-lived and immediately forgotten. Almost all of these activations do not have much sense and are transient configurations, fleeting thoughts that escape the mind without being noticed. This is usually called imagination. Interesting associations are noticed by the central executive and amplified by emotional filters that provides neurotransmitters increasing the plasticity of the circuits involved and forming new associations, pathways in the conceptual space.

Very few computational models addressing creativity have been implemented so far, the most interesting being Copycat, Metacat, and Magnificat developed in the lab of Hofstadter [25]-[26]. These models define and explore “fluid concepts”, that is concepts that are sufficiently flexible and context-sensitive to lead to automatic creative outcomes in challenging domains. Copycat architecture is based on an interplay between conceptual and perceptual activities. Concepts are implemented in a Slipnet spreading activation network, playing the role of the long-term memory, storing simple objects and abstract relations. Links have length that reflect the strength of relationships between concepts, and change dynamically under the influence of the Workspace network, representing perceptual activity in the short-term or working memory. Numerous software agents, randomly chosen from a larger population, operate in the Workspace, assembling and destroying structures on various levels. The Copycat architecture estimates “satisfaction” derived from the content of assembled structures and concepts. Relations (and therefore the meaning) of concepts and high-level perceptions emerge in this architecture as a result of a large numbers of parallel, low-level, non-deterministic elementary processes. This indeed may capture some fundamental processes of creative intelligence, although connections with real brain processes have not been explored and this approach has been applied only to design of new fonts [26].

Results of experimental and theoretical research lead to the following conclusion: creativity involves neural processes that are realized in the space of neural activities reflecting relations in some domain (in case of words knowledge about morphological structures), with two essential components: 1) distributed chaotic (fluctuating) neural activity constrained by the strength of associations between subnetworks coding different words or concepts, responsible for imagination, and 2) filtering of interesting results, amplifying certain associations, discovering partial solutions that may be useful in view of the set goals. Filtering is based on priming expectations, forming associations, arousing emotions, and in case of linguistic competence on phonological and semantic density around words that are spontaneously created (density of similar active configurations representing words).

3. Algorithms for Creation of Novel Words

In languages with rich morphological and phonological compositionality (such as Polish) novel words may appear in normal conversation (and much more frequently in poetry). Although these words are newly invented and cannot be found in any dictionary they may be understandable even without hearing them in a context. The simplest test for creative thinking in linguistic domain may be based on ingenuity of finding new words, names for products, web sites or companies that capture their characteristics. A test for creativity based on ingenuity in creating new words could measure the number of words each person has produced in a given time, and should correlate well with more demanding IQ tests.

Suppose that several keywords are given, or a short text from which such keywords may easily be extracted, priming the brain at the phonetic and semantic level. The goal is to come up with novel and interesting words that capture associations among keywords in the best possible way. Large number of transient resonant configurations of neural cell assemblies may be formed in each second, exploring the space of all possibilities that agree with internalized constraints on the phonological structure of words in a given language (phonotactics of the language). Very few of those imagined words are really interesting, but they all should sound correctly if phonological constraints are kept. Imagination is rather easy to achieve, taking keywords, finding their synonyms

to increase the pool of words, breaking words into morphemes, syllables, and combining the fragments in all possible ways.

In the brain words that use larger subnetworks common to many words have higher chance to win competition, as they lead to stronger resonance states, with microcircuits that mutually support activity of each other. This probably explains the tendency to use the same word in many meanings, and create many variants of words around the same morphemes. Creative brains are probably supported by greater imagination, spreading activation to more words associated with initial keywords, and producing faster many combinations, but also selecting most interesting results through emotional and associative filtering. Emotional filtering is quite difficult to model, but in case of words two good filters may be proposed, based on phonological and semantic plausibility. Phonological filters are quite easy to construct using second and higher-order statistics for combination of phonemes (in practice even combination of letters is acceptable). Construction of phonological neighborhood density measure requires counting the number of words that sound similar to a target word. Semantic neighborhood density measures should evaluate the number of words that have similar meaning to a target word, including similarity to morphemes that the word may be decomposed to.

Implementation of these ideas in a large scale neural model should be possible, but as a first step simplest approximations should be investigated. The algorithm should involve 3 major components:

- 1) an autoassociative memory (AM) structure, constructed for the whole language to capture its statistical properties; it stores background knowledge that is modified (primed) by keywords;
- 2) imagination may be implemented by forming new strings from combinations of substrings found in keywords used for priming, with constraints provided by the AM to select only lexically plausible strings;
- 3) final ranking of the accepted strings should simulate competition among novel words, leaving only the most interesting ones.

In the simplest case single letters are represented using temperature coding: given an alphabet of N letters each letter l is replaced by a binary vector V_l with $N-1$ zeros and 1 at the position of letter l in the alphabet. If the length of the words does not exceed K the maximum number of bits in a vector representing all letters is $n = KN$. In English there are 26 letters plus space, so $N=27$ and with $K=20$ vectors of dimension 540 are produced. A two-layer input-output network (without hidden layer) may serve as an autoassociative memory. This network is fully specified by the weight matrix \mathbf{W} . In the simplest model this is a binary correlation matrix, initially with all elements $W_{ij}=0$. Reading all words $L=(l_1l_2...l_m)$ from a dictionary and converting them to binary vectors $V_L=(V_{l_i})$ an outer product of these vectors $W_{ij}=V_iV_j$ will capture correlations between positions of all letters. The final correlation matrix is calculated as a sum of contributions over all strings L and binarized with threshold 1 or higher to filter out rare combinations of letters. For example, if there is just one word starting with aj in the dictionary then element $W_{1,37}=1$, corresponding to the first letter a (bit on at position 1) and the second letter j (bit on at position 37), but there will be many words starting with ab , so $W_{1,29}$ will be large. After binarization with threshold >1 element $W_{1,37}=0$ and $W_{1,29}=1$. Given a new string L with N_L letters the correlation matrix is used to estimate if it is lexically correct and can be accepted as a new word: if all pair-wise letter correlations that are in string L have been present in the dictionary $\mathbf{W}\tilde{N}_L V_L = V_L$ is obtained, that is the only non-negative elements of the product will be at position of 1 bits in V_L .

General properties of Auto-Correlation Matrix Memory models have been studied by Kohonen [27]. Here it is sufficient to notice that for a large dictionary correlations between pair of letters ($l_i l_{i+k}$) that are not adjacent quickly reduce to the product of probabilities of finding a given letter at position i and $i+k$ and thus all elements W_{ij} far from diagonal tend to be 1. Correlation matrices that have too many 1's accept all lexical strings and therefore are not useful. In practice correlations between letters at adjacent positions ($i,i+1$) and second neighbors ($i,i+2$) are most useful and thus the weight matrix has block structure, shown for 5 letters below:

$$\mathbf{W} = \begin{bmatrix} [D] & [x] & [x] & [0] & [0] \\ [x] & [D] & [x] & [x] & [0] \\ [x] & [x] & [D] & [x] & [x] \\ [0] & [x] & [x] & [D] & [x] \\ [0] & [0] & [x] & [x] & [D] \end{bmatrix}$$

All blocks are N by N submatrices, diagonal blocks $[D]$ for letter at position i just show that a given letter was found at this position in some dictionary words, off-diagonal blocks $[x]$ show correlations between letters at their respective positions and blocks that are further from diagonal than two units are all zero.

Although this is the simplest model it can be used to quickly filter out many spelling errors (reversing a pair of letters, as in “example” usually creates improbable strings of letters) and even identify spam messages that contain many deliberately misspelled words (lexical rules are strongly violated). Experiments with such correlation matrices show that for unrestricted dictionaries they accept too many strings (metaphorically speaking such correlation matrices do not constrain sufficiently imagination when random strings are created) and thus are not sufficient to model the process of forming phonological resonances and their associations. Weak filtering at this stage will make the final ranking of interesting novel words difficult.

There are several simple extensions to this model, at the level of word representations, more complex network models and learning algorithms. Iterative algorithms, such as Bi-directional Associative Memory model (BAM) give more accurate retrieval and guarantee greater storage capacity of associative memory. However, an interesting possibility that should be explored is to keep the algorithm as simple as possible and improve representation of words. The list of elementary units should be expanded from letters to pairs of letters, selected triplets, morphemes, or additional phonological representations, leading to an increase of the dimensionality of vectors representing words, and thus sparse correlation matrix providing stronger constraints. Heteroassociation between words and strings of sorted letters that appear in the word, or of sorted bigrams, may be used, for example associating “infinite” with “(^i)(fi)(in)(in)(it)(nf)(ni)(te)(e\$)”, where ^ and \$ signs are used for the beginning and the end of the word, and temperature coding is used to convert bigrams into binary vectors. A set of transformations based on atomic lexical elements can be applied to word converting it to quasi-phonemic transcript, or matching morpheme substrings to the word strings. This allows for simulation of particularly strong resonances corresponding to frequent phonetic elements. The representation that is used below is based on bigrams correlated with a single letter that follows them, but other representations may also be used.

To reflect constraints for filtering novel lexical strings binary weights may be replaced by correlation probabilities, taking into account word frequencies. Although such data is available for example from the British National Corpus (BNC) to reflect better contemporary vocabulary we have created our own corpus collecting about 370,000 articles from the internet (almost 1.5GB) with around 420,000 distinct word-forms. This corpus is split into a list of distinct words and the frequency of each word is counted. The correlation matrix \mathbf{W} is calculated using this list and includes frequencies of words. It is then normalized by dividing W_{ij} by the sum of all elements in a row. Other ways to normalize this matrix are also available in the program: additional position-dependent weights may stress the importance of the beginning and end atoms in words, or maximum W_{ij} value may be used instead of the sum.

In the second step various combinations of atoms should be considered. With n words and m atomic components derived from them there is $m!/(mk)!k!$ combinations with k elements to check, and even if sequential filtering is used, combining pairs first and adding more atomic components to highly probable combinations only, the number of such combinations will grow very quickly. Words are always created in some context. In practical applications we are interested in creating novel names for some products, companies or web sites. Reading descriptions of such objects people pick up important keywords and their brains are primed, increasing probability of creating words based on atomic components found in keywords and additional words that are strongly associated with these keywords. The key to get interesting new words is to supply the algorithm with a broad set of priming words somehow related to the main concept. In our model this is realized by priming with enhanced set of keywords generated from Wordnet synsets (sets of synonyms) [28] to original keywords. The extended set of keywords may then be checked against the list generated from our corpus to get their frequencies. To account for priming the main weight matrix is modified by adding $\mathbf{W} + \mathbf{W}_p$, where \mathbf{W}_p is the weight matrix constructed only from the keywords. \mathbf{W}_p is multiplied by a factor α that controls the strength of the priming effect. Using very large α makes the background knowledge contained in the weight matrix \mathbf{W} almost irrelevant; the results are limited only to a few words because the program filters out almost all words as the priming set is not sufficient to learn acceptable correlations. A binary \mathbf{W}_p matrix may also be used if each row of the combined matrix is divided by its maximum element.

In the brain priming of some words leads to inhibition of others. This may be simulated by implementation of negative or inhibitory priming that decreases the weights for words that are antonyms of keywords. For example, while creating words from such keywords as unlimited happiness combinations of “unhappy” types, although formally interesting, should be avoided and usually do not come to our mind. The algorithm for creating words works at the syntactic level and does not try to analyze the meaning of the words. Negative priming may also partially help to avoid controversial words, for example with obscene content or unwanted side-effect meanings. In the simple version of the algorithm considered here meanings are not explicitly coded, although using enhanced sets of keywords helps to capture semantics to some degree. Two desired characteristics of a software product described by keywords “powerful” and “boundless”, analyzed at the

morpheme level will lead to a perfect word “powerless” with a very high score, yet in most cases this association will not come to mind of people, inhibited at the semantic level. This score could be lowered by negative priming, but in current implementation of our algorithm such words get lower ranking only in the final stage, when the “relevance” and “interestingness” filter is applied, and associations of the created word are searched for. If strong associations with some antonyms of keywords are discovered the word gets low ranking. Novel words should not have too much resemblance with the words that are already in the dictionary because they will be not be treated as new, only as misspelled words. A set of words should be rather diverse, otherwise even interesting new words may be perceived as boring, too many variants on the same theme. On the other hand there should be some similarity to keywords – without similarity it would impossible to form desired associations. One way to estimate how interesting the word may seem to be is to evaluate its “semantic density”, or the number of potential associations with commonly known words. This may be done by calculating how many substrings within the novel word are lexical tokens or morphemes. For longer morphemes similarity rather than string equivalence should be sufficient. If several substrings are similar to morphemes or words in the dictionary the word will certainly elicit a strong response from the brain networks and thus will be seen as interesting.

The influence of subjective, personal bias can also have an impact when judging the obtained results. It may be at phonological or semantic level, related to some idiosyncratic preference that cannot be found in any dictionary. This aspect is very easy to miss in subjective evaluation of a given word. Knowing individual preferences and favorite expressions the algorithm could be to some degree personalized.

The normalized correlation matrix combined with the priming set matrix is and renormalized is used to evaluate the new words that are created by combining all atoms derived from extended keywords that form words of a given length. Similarity between words is calculated using edit distance for atomic elements, with no penalty if all atoms are identical, and +1 penalty for deletion, substitution or insertion of a single atom. To avoid problems with different number of atoms for words $L=(a_1a_2a_3a_4\dots a_w)$ of different length the word is treated as a closed cycle, with $a_{w+1}=a_1$, $a_{w+2}=a_2$ and so on, with a_N being always the last atom. For every new word four test scores are calculated: the total product of all ngram frequencies, the minimal ngram frequency (to eliminate weak links), the longest common subsequence and the longest common substring with words in the priming set (without any transformation).

Results presented below have been created with the algorithm presented above and the following additional requirements: 1) new words should differ at least by two atoms from words from the dictionary; 2) new words should differ at least by three atoms from each other; 3) the length of the generated words should not exceed seven atoms; 4) the longest common substring has at least two atoms; 5) if not stated otherwise, position is included within each ngram.

4. Results

Two simple examples are presented here. First interesting names for a website offering “shoes” is searched for. From the company brochure the priming set of keywords is extracted, consisting of such words as “running, sport, youth, health, freedom, air”. Results of 4 algorithms that use different word representations are presented below.

For an extended ngram model results are ordered by product of all ngram probabilities. Among the best 20 words several interesting words appear: (4) shoebie, (11) airenet, (12) runnyne, (20) sportip. The number in the brackets is the position on the list.

Using extended sorted ngrams removes some words from the list (shoebie, sportip) and adds new ones: (2) airenet, (5) windway, (14) funkine, (15) runnyne and (19) moveman.

The third model is based on reversed, extended ngrams, removing airenet, runnyne, sportip, windway, moveman, and adding: (2) funkine, (3) shoebie, (8) runably, (18) sporist.

When both straight and reversed extended ngrams are used the outcomes are as follows: removed airenet, runnyne, windway, funkine, moveman, sporist, and added: (1) shoebie, (6) winaway, (9) runably, (16) sportip and (18) runniess. Some mutations of previous words appeared here (winaway, runnies) that are quite similar to previous words, which leads to an idea to group first variations of one word and rank not single words but whole groups. This however could be computationally more complex.

Google search for these words shows that some of them have been already invented by people, although not necessarily applied in the context of shoes. For example “airnet” is a great name for wireless services, and “Winaway” is a name of a racing greyhound champion. Table I summarizes the results, quoting approximate number of entries in Google search engine at the end of the year 2006. Although these words are relatively rare most of them have been already used in various ways. The domain www.sportip.com was for sale for \$17,000.

Table I: Summary of Interesting Words Related to “shoes”

Word	# in Google	Remarks
airenet	770	Mostly wireless networks
funkine	70	Music term, “Funk in E”
moveman	24000	Mostly moving companies
runably	New	
runniess	New	
runnyme	220	runnyme.de, company name
shoebie	2700	Slang word, many meanings
sporist	16400	sporist.com, used in Turkish language
sportip	2500	Web sites, in many languages
winaway	2400	Dogs, horses, city name
windway	99500	windway.org, popular, many meanings

Table II: Summary of interesting words related to “creative ideas”

Word	# in Google	Remarks
crealin	970	Used in many languages
ideates	690	ideates.com, defined in some dictionaries
invelin	10	Rare biomedical term
inveney	New	
smartne	16000	CISCO product
talenev	20	Rare female name
timepie	610	Name for watches
visionet	96100	Popular name for many products and services

The second example came from a real request for finding a good company and portal name; the company wanted to stress creative ideas, and the priming set consisted of such concepts as idea, creativity, portal, invention, imagination, time, space. The 4 algorithms applied in this case gave:

- extended ngrams: (6) ideates, (7) smartne, (9) inveney, (13) timepie,
- extended sorted ngram: (1) ideates, (8) smartne, (10) inveney, (11) timepie, (2) talenev,
- extended reversed ngrams: (1) ideates, (3) talenev, (6) inveney, (12) crealin, (18) invelin, (19) visionet, and removed smartne, timepie
- extended sorted and reversed ngrams: (2) ideates, (9) inveney, (10) talenev, (14) crealin, (18) visionet; removed smartne, timepie, invelin.

It is possible to change the final filtering criteria but since all score systems are strongly related to each other the results differ only by small variations in the words and are of similar quality. In another experiment starting from an extended list of keywords: “portal, imagination, creativity, journey, discovery, travel, time, space, infinite”, more interesting words has been generated, with about $\frac{3}{4}$ already used as company or domain names. For example, creatival is used by creatival.com, creativity is used by creativity.com. Some words have been used only a few times (according to the Google search engine), for example discoverity that can be derived from: disc, disco, discover, verity, discovery, creativity, verity, and may mean discovery of something true (verity). Another new interesting word is digventure, because it is easy to pronounce, and both “dig” and “venture” have many meanings and thus many associations, creating a subnetwork of activity in the brain that resonates for a long time.

5. Discussion and conclusions

Neurocognitive approach to the creation of novel words in the brain presented here is obviously quite speculative and the actual implementation is still rather simplistic, but it seems to open the door to modeling of creativity in this specialized domain. Brain imaging and electrophysiological studies of brain activity during invention of new words, as well as during analysis of novel words, would make an interesting test of neurocognitive approach to creativity and may be done with methods already used to study word representations

[6][7]. Probing associations and transition probabilities between brain states using priming techniques [23] should help to understand better associations that may be used to extend keyword lists. Research program on creativity that includes neuroscience, cognitive psychology and theoretical modeling, focused on word representation and creation, could be an entry to a detailed understanding of this fascinating brain processes.

Few examples of results presented support the idea that creative processes are based on ordinary cognitive processes and that understanding creativity and developing computational models of creativity in some areas may actually be feasible. Creativity requires prior knowledge, imagination and filtering of the results. Imagination should be constrained by probabilities of composition of elementary operations, corresponding to activations of specific brain subnetworks. Products of imagination should be ranked and filtered in a domain-specific way. The same principles should apply to creativity in design, mathematics, and other domains, although in visual or abstract domain elementary operations and constraints on their compositions are not so easy to define as in the lexical domain. In arts emotional reactions and human reactions to beauty are rather difficult to formalize. Nevertheless it should be possible to create a network that learns individual preferences evaluating similarity to what has been assessed as interesting. It is sufficient to observe how long and where a person looks in the art gallery to learn preferences and to create a new painting that would fit this person's taste. In abstract domains various measures of relevance or interestingness may be used for filtering, but to be interesting creative abstract designs (for example in mathematics) will require rich conceptual space, reflecting many neural configurations that may be potentially active.

The algorithms used here to create new words are quite efficient and may be used in practice to generate interesting names for products, companies, web sites or names of various entities. To estimate practical usefulness of such algorithms their results should be compared with human inventiveness in a larger number of cases. Humans can obviously evaluate results in a better way than our scoring system. It should be quite interesting to see how word creativity tests correlate with more sophisticated and well established tests. Computational models of creativity outlined in this paper may be implemented at a different level of neurobiological approximations, from detailed neural models to simple statistical approaches. Even simple algorithms are capable of producing interesting words, and the fact that many of these words have already been invented by humans shows that these algorithms are able to abstract some important properties of the creative process. Unfortunately they also leave many words that should be manually removed from the final list. The drawback of the current implementation were already mentioned and some improvements proposed. To correct them several steps should be taken, the most important being:

- learning words using phonetic instead of lexical dictionaries;
- using quasi-morpheme decomposition of words, rather than lexical atomic units;
- introduce negative priming;
- add learning from user judgments.

Learning may be introduced in many ways, the simplest is to adapt some global parameters, such as the priming and the inhibition factors, adding weights to different morphemes depending on their frequency etc. Self-organizing mapping algorithms may also be used to create word representation from the subatomic components. In this paper only the simplest correlation-based approach has been investigated, but it is clear that many possibilities remain to be explored and eventually a more detailed models of word associations at the atomic level will be created.

Acknowledgment

Support by the Polish Committee for Scientific Research, research grant 2005-2007, is gratefully acknowledged.

References

- [1] R. Wilson and F. Keil, Eds. *MIT Encyclopedia of Cognitive Sciences*, MIT Press, 1999.
- [2] M. Runco and S. Pritzke, Eds. *Encyclopedia of creativity*, vol. 1-2, Elsevier, 2005.
- [3] R.J. Sternberg, Ed. *Handbook of Human Creativity*. Cambridge: Cambridge University Press, 1998.
- [4] F. Pulvermüller, *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge, UK: Cambridge University Press, 2003.
- [5] F. Pulvermüller, Brain reflections of words and their meaning. *Trends in Cognitive Sciences* Vol. 5, pp. 517-524, 2001.

- [6] F. Pulvermüller, A brain perspective on language mechanisms: from discrete neuronal ensembles to serial order. *Progress in Neurobiology* vol. 67, pp. 85-111, 2002.
- [7] F. Pulvermüller, Y. Shtyrov and R. Ilmoniemi, Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience* Vol. 17(6), pp. 884-892, 2005.
- [8] W. Duch, Platonic model of mind as an approximation to neurodynamics. In: *Brain-like computing and intelligent information systems*, ed. S-i. Amari, N. Kasabov, Springer, Singapore, Chap. 20, pp. 491-512, 1997.
- [9] W. Duch, Categorization, Prototype Theory and Neural Dynamics. *Proc. of the 4th International Conference on Soft Computing'96*, Iizuka, Japan, ed. T. Yamakawa and G. Matsumoto, pp. 482-485, 1996.
- [10] A. Jensen, *The g Factor: The Science of Mental Ability*. Westport CT, Praeger, 1998.
- [11] D.L. Molfese, Predicting Dyslexia at 8 Years of Age Using Neonatal Brain Responses. *Brain and Language* vol. 72, pp. 238-245, 2000.
- [12] D.L. Molfese, and V.J. Molfese, The Continuum Of Language Development During Infancy and Early Childhood: Electrophysiological Correlates. In C. Rovee-Collier, L. P. Lipsitt, and H. Hayne (Eds.), *Progress in Infancy Research*, Vol. 1. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 251-287, 2000.
- [13] A.D. Baddeley, Is Working Memory Still Working? *European Psychologist* vol. 7, pp. 85-97, 2002.
- [14] N. Cowan, *Working memory capacity*. New York, NY: Psychology Press 2005.
- [15] D.S. Ruchkin, J. Grafman, K. Cameron, and R.S. Berndt, Working Memory Retention Systems: A State of Activated Long-Term Memory, *Behavioral and Brain Sciences* Vol. 26(6), pp. 709-728, 2003.
- [16] D. Caplan, G.S. Waters, Verbal Working Memory and Sentence Comprehension. *Behavioral and Brain Sciences* vol. 22, pp. 77-94, 1999.
- [17] H. Damasio, T.J. Grabowski, D. Tranel, R.D. Hichwa and A.R. Damasio, A neural basis for lexical retrieval. *Nature* vol. 380, pp. 499-505, 1996.
- [18] A. Martin, C.L. Wiggs, L.G. Ungerleider and J.V. Haxby, Neural correlates of category-specific knowledge. *Nature* 379, 649-652, 1996.
- [19] L.L. Lapointe, *Aphasia And Related Neurogenic Language Disorders*. 3rd Ed, New York, NY: Thieme, 2005.
- [20] E.K. Vogel, AW McCollough, MG Machizawa, "Neural measures reveal individual differences in controlling access to working memory". *Nature* vol. 438, pp. 500-503, 2005.
- [21] S. Grossberg, Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31, 423-445, 2003.
- [22] S.A. Mednick, The associative basis of the creative process. *Psychological Review*, 69, 220-232, 1962.
- [23] A. Gruszka, E. Nęcka, Priming and Acceptance of Close and Remote Associations by Creative and Less Creative People. *Creativity Research Journal* Vol. 14(2), pp. 193-205, 2002.
- [24] T. Wellens, V. Shatokhin, and A. Buchleitner, Stochastic resonance. *Reports on Progress in Physics* Vol. 67, pp. 45-105, 2004.
- [25] D.R. Hofstadter, *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. NY: Basic Books, 1995.
- [26] J. Rehling, *Letter Spirit (Part Two): Modeling Creativity in a Visual Domain*. PhD Thesis, Indiana University, 2001.
- [27] T. Kohonen, Correlation matrix memories, *IEEE Transactions on Computers C*, Vol. 21, pp. 353-359, 1972.
- [28] Wordnet, available at <http://wordnet.princeton.edu>
- [29] W. Duch, Brain-inspired conscious computing architecture. *Journal of Mind and Behavior*, 26(1-2), 1-22, 2005.



Włodzisław Duch graduated from the Nicolaus Copernicus University in 1977 and currently is heading the Department of Informatics at this university and is a visiting professor at the Nanyang Technological University in Singapore. His research interest includes computational models of brain functions and computational intelligence. (To find his home page Google: Duch)



Maciej Pilichowski obtained his MSc from the Nicolaus Copernicus University in 1998 and currently is a graduate student.