

Intuition, Insight, Imagination and Creativity

Włodzisław Duch

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

Google: Duch

Abstract—Can computers have intuition, insights and be creative? Neurocognitive models inspired by the putative processes in the brain show that these mysterious features are a consequence of information processing in complex networks. Intuition is manifested in categorization based on evaluation of similarity, when decision borders are too complex to be reduced to logical rules. It is also manifested in heuristic reasoning based on partial observations, where network activity selects only those path that may lead to solution, excluding all bad moves. Insight results from reasoning at the higher, non-verbal level of abstraction that comes from involvement of the right hemisphere networks forming large “linguistic receptive fields”. Three factors are essential for creativity in invention of novel words: knowledge of word morphology captured in network connections, imagination constrained by this knowledge, and filtering of results that selects most interesting novel words. These principles have been implemented using a simple correlation-based algorithm for auto-associative memory. Results are surprisingly similar to those created by humans.

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

One of the objections against computational intelligence considered by Alan Turing in his famous article “Computing machinery and intelligence” [1] recalls Lady Lovelace's objection (written in her memoirs in 1842) that a machine can “never do anything really new”, and in particular the Analytical Engine of Babbage (an early idea for universal computer) “has no pretensions to originate anything. It can do *whatever we know how to order it to perform*”. Turing’s response can be summarized as: “the evidence available to Lady Lovelace did not encourage her to believe” that machines could be creative, although “It is quite possible that the machines in question had in a sense got this property” because “suppose that some discrete-state machine has the property. ... universal digital computer ... could by suitable programming be made to mimic the machine in question”. It is difficult to ascertain that something is really new, and Turing admits that “Machines take me by surprise with great frequency”.

The last section of Turing’s article is devoted to learning machines as our best hope to realize computational intelligence and creativity. After proposing (albeit in a very vague terms) “the child machine” in the final paragraph of the paper the author writes: “We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best.” This has indeed proved to be true and before the turn of the century – as Turing predicted – computers exceeded human level competence in chess. However, the connection between memory capacity and speed of calculations in chess is quite obvious, therefore the famous Big Blue – Kasparov

match has been accepted more as a demonstration of sheer computer power rather than true machine intelligence. Turing suggested also another approach: “It can also be maintained that it is best to provide the machine with the best sense

“We may hope that machines will eventually compete with men in all purely intellectual fields”, wrote Alan Turing. He believed that learning machines can be creative, and proposed to develop both computer chess and a “child machine”, or embodied intelligence approach.

organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child. Things would be pointed out and named, etc.” Many people turn now to this, much harder, approach hoping that autonomous mental development using real embodiment of perception/action in robot brains may be the answer (see special issue of CIM [2]). Highly abstract symbolic activity and fully embedded processes drawing on perception and exploration of the world are two extremes, with a lot of fertile ground in between.

Many low-level cognitive functions [3] involving perception and motor control have already reasonable neural models that capture more details every year. With growing sophistication of algorithms, software implementations, and new inspirations from neuroscience the field seems to be on a good track and some notable successes are already evident [4], although reaching animal-level proficiency may still take some time. Understanding and modeling of higher cognitive functions, including visual and auditory scenes, the use of language, thinking, reasoning, planning, problem solving or building architectures to coordinate all cognitive functions is in much worse shape. Models of thinking processes have been dominated mostly by search and rule-based symbolic Artificial Intelligence (AI) algorithms, with a few toy examples based on connectionist approaches in linguistic domain [2]. Consciousness is considered to be the most mysterious of all mental phenomena [5], but it may not so difficult to realize in artificial systems. Brain-like information processing have to lead to claims of consciousness in systems that are able to comment on their internal states [6]. These comments are needed to make sense of narrative history of one's own life as well as to learn any complex skill that requires coordination of perceptions, reasoning and actions [6].

Consciousness should arise in systems based on brain-like information processing that may comment on their own internal states.

Unraveling detailed brain circuits involved in creation of such comments, as well as computational implementation of complex systems based on recurrent modules that implement this type of information processing, may take long time. Research on many aspects of consciousness is quite active [5], but there are other, quite neglected faculties that all child machines or any artificial minds must possess.

Arguably, three most important (and most mysterious) faculties of the mind needed for intelligent behavior are intuition, imagination and creativity. Babies and animals do not reason making logical inferences, be it crisp or fuzzy, but even birds use intuition, imagination and creativity to solve problems [7]. Computers need to show similar qualities.

Intuition

The MIT Encyclopedia of Cognitive Sciences [8] has 10 articles devoted to various aspects of logic and almost 100 index references to logic. Intuition is not mentioned in the index at all, although several articles mention definitions that agree with some intuitions. The word "intuitive" in biology, psychology, mathematics, physics and sociology is treated as synonym of naive understanding in these fields. Yet in everyday activity very few people (and certainly no animals) base their decisions on logical analysis of all options. Most cognitive functions, such as understanding of human and animal intentions and emotional states, meaning of words or creative thinking, cannot be reduced to logical operations. Why then intuition is played down and so much effort is spent on logic? Perhaps we have been blinded by the apparent power of logics in the early models of brain functions, leading to AI focus on logical methods in symbol manipulation for problem solving. Computational functionalism in philosophy of mind separated neural and mental processes focusing on symbolic analysis of thinking processes. Logical approaches to truth, language and understanding of behavior gave rise to many technical questions keeping experts busy for many years, although little progress towards the initial goal has been made. It is much easier to develop existing theories and formalisms rather than to come up with new conceptualization of the problem.

Intuition is defined in dictionaries as immediate knowing without the conscious use of reasoning, or cognition without evident rational thought and inference. Deliberate thinking is critical, analytic and reasoning-like, while intuitive thinking is rapid, effortless, and perception-like. The subject of intuition has been abandoned by science, left mostly to esoteric psychology or at best to psychoanalysis. Only recently scientific psychology showed some interest in intuition. Social cognitive neuroscience views implicit learning processes as the cognitive substrate of social intuition [9]. After publication of the book "Intuition: Its Powers and Perils" by D.G. Myers [10] review in Scientific American called intuition "... a rich emerging field of scientific inquiry" (quoted from the book cover). In experimental psychology studies of subliminal priming, implicit memory, automatic processing, emotional cues, non-verbal communication, prejudices and stereotypes, subconscious use of heuristics, decision making, blindsight and other brain damage phenomena, are all relevant to understanding intuition. Obviously simple perception leading to object recognition in any sensory modality, does not require logical reasoning but brings immediate knowing of the objects seen, heard or touched. Psychologists and neuropsychologists have thus given the research on intuition some respect, although it still lacks multidisciplinary focus.

For many years intuition has been neglected but now it is "... a rich emerging field of scientific inquiry", according to review in Scientific American.

From computational perspective modeling intuition is relatively simple. Decisions of neural networks (or other models) that learn from data frequently cannot be justified in terms of logical rules. In some cases logical rules that have similar or even higher predictive power may be extracted from trained neural networks [11]. In other cases judgments based on overall similarity provide better decisions. For example, data that is generated from a single oblique Gaussian probability density function will be classified with high accuracy using a single reference vector \mathbf{R} with Mahalanobis metric $\|\mathbf{X}-\mathbf{R}\|$ that measures dissimilarity between the query and the reference vector. Neural networks may easily learn this type of similarity evaluation, but there is no simple way to express equivalent knowledge in terms of logical rules. Only for additive metric functions $\|\mathbf{X}-\mathbf{R}\| = \sum_{i=1} W_i d(X_i, R_i)$, where $d(\cdot)$ evaluates dissimilarity for feature X_i , fuzzy interpretation in terms of membership functions is possible. Using product norm and exponential transformation:

$$T(\mathbf{X}, \mathbf{R}) = \prod_{i=1} \mu_i(X_i), \text{ with } \mu_i(X_i) = \exp(-W_i d(X_i, R_i)),$$

identical decision borders as with the prototype-based rules may be recreated. For that reason it has been conjectured [12] that prototype-based rules (P-rules) in threshold or nearest-neighbor form are more general than fuzzy or crisp rules (F-rules, C-rules), offering more flexibility and biological plausibility in modeling of perception and decision making (see [13] on interesting relations between uncertainty in data, multilayer perceptrons and fuzzy rules). If several such prototypes are needed P-rules can still handle the problem in an easy way while approximations based on fuzzy rules will almost always be of poor accuracy and will require many rules, making the whole system incomprehensible. Psychologists have noticed that rules and similarity judgments form a continuum, with logical rules (including threshold logic and fuzzy logic rules) applicable in relatively simple cases, while prototype-based rules applicable in situations when many factors are simultaneously taken into account for similarity judgments [14]. For example, medical doctors may use simple norms based on thresholds for some tests, but in case of emergency they have to make fast intuitive judgments, taking many factors into account. Experience leads to intuition, and it is obviously related to similarity evaluation and memorization of many prototypes. Even for simple benchmark medical data, a single P-rule may offer more accurate explanation than sets of logical rules [15].

Prototype-based rules offer a good model of recognition-based intuition, in many cases more accurate and simpler to comprehend than fuzzy rules.

Intuition is usually invoked in context of reasoning and decision making. Herbert Simon claimed that AI has reached the stage where intuition, inspiration and insight could be modeled [16]. Intuition in problem solving has two defining characteristics: (1) the solution has to be reached rapidly, and (2) explanation why the steps leading to solution has been selected could not be given. In various experiments novices and experts solving the same problem were compared and the use of intuition has been clearly correlated with the ability to evaluate similarity and with the number of patterns stored in long term memory. Knowledge obtained through implicit learning or derived from partial observations (in contrast to the usual supervised learning situation, when full knowledge is provided) over a long period of time cannot be used directly in explicit reasoning. It is represented in a diffuse, rather weak connections, partially in the right brain hemisphere, and thus cannot be accurately summarized in symbolic form. Some attempts to capture intuition in chess have been made recently [17], using rather sophisticated representational scheme. The claim is that “the postulated architecture models chess intuition as an emergent mixture of simultaneous distance estimations, chunk perceptions, abstract role awareness, and intention activations.” Our brains constantly learn to pay attention to relevant features and remember many patterns. Even in tasks for which rules of correct actions exist intuitive learning comes before rules are discovered.

Knowledge required for solving pattern recognition problems is usually quite limited, in most cases gained from a single dataset given for training. Problems that require systematic reasoning are solved in AI by using a lot of background knowledge, selecting and combining relevant rules in a process of searching for solution. Combinatorial explosion may be avoided if high level macro-operators are used as a shortcut; such strategies are based on the idea of chunking (grouping) knowledge in hierarchical fashion, used in some AI systems such as SOAR [18]. Hierarchical Temporal Memory model, recently proposed as a general cortex mechanism, works using this principle [19], learning common spatial and temporal sequences to discover causes. This is quite similar to hierarchical correlation learning used already in the pandemonium model [20] almost half a century ago, used also in brain-inspired vision systems [4].

Intuitive machines should learn from partial observations, correlating subsets of features to create chunks of knowledge. In many domains strong symbolic rules are not known. Instead, implicit learning creates a number of neural modules that capture some correlations between selected variables. This is quite common in natural situations, for example observing animal behavior patterns various cues are memorized, and predictions of future

activity and intentions are made. Some animals actively tease predators to test their reactions and gain valuable knowledge [7]. In the original PDP Books [21] several articles have been concerned with problems that required combinatorial constraint satisfaction. Relations between two or three variables constraining their possible values have been defined, and Boltzmann machines and harmony theory have been used to search for self-consistent states of these networks. Such methods proved to be rather inefficient because stochastic training algorithm does not scale well with the size of the problem. Recently multi-layer restricted Boltzmann machines and deep belief networks have been introduced [22], based on stochastic algorithms and binary representations, but their use was restricted to pattern recognition problems so far.

Solutions of complex problems, including inferences about observed behavior, combine systematic search with intuitive recognition based on partial observations. An approach to capture essential aspects of intuitive reasoning based on systematic search has been proposed in [23]. Intelligence is sometimes a matter of fast intuitive estimation of what can, and what cannot be true. Suppose that a number of relations between small subsets of all features characterizing complex system are known *a priori* or are derived from observations. For example 3 features may be constrained by some function $F(A,B,C)$, logical relations, or by observation that (A,B,C) may take only restricted values. All basic laws of physics have this form. Relations may also be found for changes in feature values. In the simplest case one may assume $\Delta A=0$ for no change, $\Delta A = +$ for increase, and $\Delta A = -$ for decrease. The speed of changes may of course be quantized into more steps. If only 3 values are admitted, for 3 variables there are $3^3=27$ possibilities, from all variables decreasing: $(\Delta A,\Delta B,\Delta C) = (-,-,-)$, to all variables increasing: $(\Delta A,\Delta B,\Delta C) = (+,+,+)$. Introducing $A=F(B,C)$ relation that is either additive $A=B+C$, multiplicative $A=B \cdot C$ or inverse additive $A^{-1}=B^{-1}+C^{-1}$ (most laws of physics are in this form) excludes 14 out of the 27 possible patterns of $(\Delta A,\Delta B,\Delta C)$ triples, for example $\Delta A=0$ (constant) is impossible if both ΔB and ΔC decrease or if both increase). It is quite surprising that when it comes to change many relations show qualitatively the same behavior, shown in Fig. 1 for (V,I,R) variables for $V=IR$ (Ohm's law). There are 13 true facts and 14 false ones, with the strength of true relations being greater for $(\Delta V,\Delta I,\Delta R) = (+,+,+)$ then for $(\Delta V,\Delta I,\Delta R) = (+,+,-)$, as the first one is always true and the second one depends on relative speed of ΔI and ΔR changes. Note that averaging over all observations will show no correlations between ΔV and $\Delta I,\Delta R$, as all three situations: $(+,+,-)$, $(0,+,-)$, $(-,+,-)$ are possible. Instead of calculating correlations facts are remembered and the response of a node is getting stronger with growing number of observations, as illustrated in Fig. 1 using different sizes of gray balls. This function is all that is needed for qualitative reasoning; it may be represented by:

$$F(X)=F(A,B,C)=\exp(-\beta \|X-(-1,-1,-1)\|^2)+ \dots \exp(-\beta \|X-0\|^2)+ \dots + \exp(-\beta \|X-(+1,+1,+1)\|^2)$$

with a large constant β .

It is quite likely that our knowledge of qualitative physics is internalized in such simple manner; if the predator runs quickly the distance decreases fast and the time left before deadly encounter is short, so qualitative relations between time, speed, and distance are important. Checking if something is possible does not require writing and solving equations; if the response of neural node $F(A,B,C)>0$ than relation between (A,B,C) features is not violated. A soft penalty function $F(A,B,C) = \exp[-\beta(A-f(B,C))^2]$ for violation of $A=f(B,C)$ relation may be used if real feature values instead of changes are preferred. If the $A=f(B,C)$ relation represents a law of nature β may be estimated from the accuracy of A, B, C measurements, if this is just a preference relation value of β may be selected to account for it. Such mechanism allows to say what is possible and what is unlikely in purely intuitive way. Impossible patterns of feature values simple "do not come to mind", as there is no activation that corresponds to them.

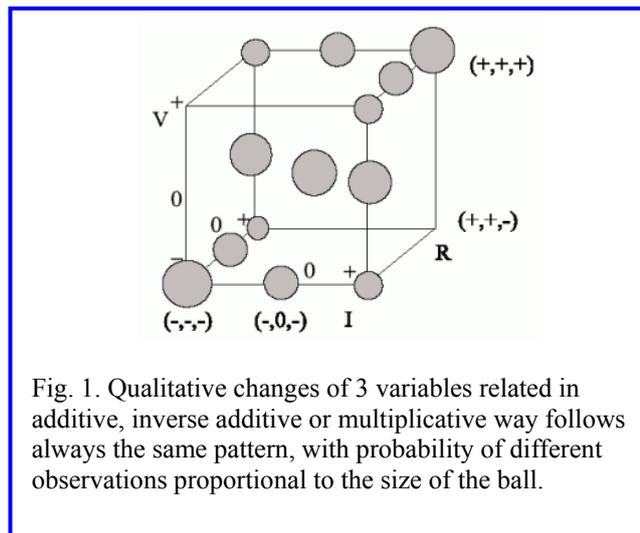


Fig. 1. Qualitative changes of 3 variables related in additive, inverse additive or multiplicative way follows always the same pattern, with probability of different observations proportional to the size of the ball.

Another surprising fact is that in complex situations expectations generated using such weak constraints – about half of the relations being true, and the other half false – are very useful. If many relations are applicable for N ternary features, out of 3^N possible combinations of possible values only a few will be in agreement with all constraints that restrict the kind of situations that may really happen. For example, the following relations:

$$f(A_1,A_2)=A_3; f(A_2,A_3)=A_4; \dots f(A_{N-2},A_{N-1})=A_N$$

leave only $4N+1$ solutions that agree with all constraints, a negligible fraction of all 3^N patterns. Such knowledge based on partial observations may be implemented in several ways [23]. A network of “knowledge atoms” containing $F(A_i, A_{i+1}, A_{i+2})$ relations that represent correlations among a subset of variables (they may be discovered in data using algorithms similar to association rules mining) may be arranged in one-dimensional array, connected to relevant input features. If the values of any two variables in the node (A_i, A_{i+1}, A_{i+2}) are known then this nodes may provide unique value (or at least some constraints) for the third variable. Search with at most $N-2$ steps, in each step selecting nodes that have only one unknown variable, determines all missing values. Slightly more difficult situation occurs when only one feature in each node has specific value, for example A_1 and A_4 . This requires systematic reasoning: suppose that A_2 has some specific value, is that possible in view of all known constraints and fixed values of variables? Again all that is needed is to check whether $F(X)>0$ for subsets of features with known values. If only A_1 and A_4 are known assume that A_2 is either $-$, 0 , or $+$, starting 3 branches of a search tree. In the first step relation $f(A_1, A_2)=A_3$ determines A_3 , in the second step $f(A_2, A_3)=A_4$ is checked for all 3 branches, stopping the search is both relations are not fulfilled. A useful heuristics is to look for maximally constrained feature, that is to find first the feature that may assume only one possible value; this requires checking if $F(X)=0$ for all other values. Fixing the values of successive features restricts the remaining features, making the search process in most cases rather trivial.

In the PDP book [21] a simple electric circuit with battery and two resistors has been analyzed using Boltzmann machines and harmony theory. The circuit (Fig. 2) can be fully described using 7 variables: current I , 3 voltages V_i and 3 resistances R_j . Most students of physics or electrical engineering will answer questions such as: if R_2 increases, and R_1 and V_t are kept constant, what will happen to the current I and how will V_1, V_2 change? Although a novice may try to deduce the answer transforming Ohm’s and Kirchoff’s equations to calculate I, V_1, V_2 from known values, an expert will answer intuitively without any deliberation. If the question will change the novice will again have to solve equations, while the expert will come up intuitively with immediate answer.

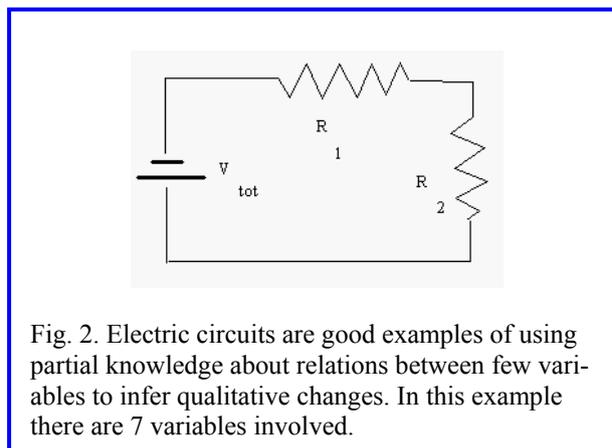


Fig. 2. Electric circuits are good examples of using partial knowledge about relations between few variables to infer qualitative changes. In this example there are 7 variables involved.

What useful knowledge do we have here? Both the novice and the expert know Ohm’s law $V=IR$ and know that $V_t=V_1+V_2$, but only in the brain of an expert, through frequent observations how currents and voltages change in real circuits, qualitative behavior captured in the cube (Fig. 1) has been internalized. Focusing on all elements 5 applicable laws are noticed: $V_t=IR_t, V_1=IR_1, V_2=IR_2, R_t=R_1+R_2, V_t=V_1+V_2$. Thus the total heuristic function is a product of 5 identical factors:

$$F(\mathbf{X}) = F(V_t, V_1, V_2, R, R_1, R_2, I) = f(\Delta V_t, \Delta I, \Delta R_t) f(\Delta V_1, \Delta I, \Delta R_1) f(\Delta V_2, \Delta I, \Delta R_2) f(\Delta R_t, \Delta R_1, \Delta R_2) f(\Delta V_t, \Delta V_1, \Delta V_2)$$

There are $3^7 = 2187$ different 7-dimensional ternary vectors \mathbf{X} , but only for 111 of them give $F(\mathbf{X})>0$, other values lead to one or more factors equal to zero. Knowing that $\Delta V_t=0, \Delta R_1=0$, and $\Delta R_2=+$ the changes of the four remaining variables should be found. It is easy to check that assuming $\Delta V_1=0$, or $+1$, or -1 , does not zero $F(\mathbf{X})$, as the unknown change in current I and the voltage V_2 may be consistent with any change in V_1 . However, $\Delta R_t = +$ is the only solution as $f(\Delta R_t, \Delta R_1=0, \Delta R_2=+)$ is 0 in other cases. The current I has to decrease now, and this leads to decrease of V_1 and increase of V_2 . No equations are transformed to solve for unknown values, only the response of functions that relate unknown to known variables is checked, in the first pass finding those features for which some factors may be uniquely determined, and if this is not possible assuming finding the most constrained feature and creating search tree with several branches.

Networks of knowledge atoms may solve many problems where partial observations lead to some constraints, facilitating intuitive reasoning. Is this really the mechanism behind intuitive problem solving? A number of testable predictions on human intuitive performance can be generated assuming this mechanism. For example, learning only theory does not lead to good intuitions, observations of how things change are needed. Good car drivers may have problems recalling driving rules, they just make correct assumptions and predictions. If the problem admits more than one solution how likely is it that a student will find all solutions? This should depend on the working memory load, or complexity of the search, needed to find all solutions. In complex situations hierarchical decomposition of the problem is necessary, depending on the questions asked. For example, ele-

ments of complex electrical circuits may be decomposed into larger blocks, there is no need to assign values to all variables. People in such cases analyze graphical structure of connections and nodes representing the problem, starting from elements mentioned in the problem statement.

We have created several software implementations of algorithms for learning from partial observations that quickly find all solutions if many discrete feature values are missing (T. Maszczyk, J. Rzepecki, W. Duch, in preparation). Problems of this type are somewhere in between pattern recognition and symbolic reasoning problems. Neural networks may be used as heuristics to constrain search processes (a core AI technology) in problem solving. Robots, including autonomous vehicles, need to combine reasoning with pattern recognition in a real time. Intuitive evaluation of possible solutions to global goals may help to generate rough plans, find optimal patterns for behavior of a robot. Other applications include games as well as industrial installations, where operators learn to interpret complex signaling patterns. Collecting data for challenging problems of this kind would be very worthwhile, encouraging the development of more algorithms to solve them.

Insight

Intuition and insight have some similarities, but the sudden Aha! experience that accompanies solutions of some problems has distinct character [24]. Insight is usually preceded by an impasse, frustration after a period of lack of progress, followed by conviction of the imminence of solution, frequently after a period of incubation when the problem is set aside. New way of looking at the problem that leads to the solution is accompanied by a great excitement and understanding. The mild version of Aha! experience is fairly common during discussions when difficult concept or a confusing description of some situation is finally grasped. Herbert Simon believed that the EPAM (Elementary Perceiver And Memorizer) model developed by Feigenbaum and himself in the early sixties [25], combined with his GPS (General Problem Solver) model [26], explains insight. The initial process of searching for the solution reaches dead end, but during the search new features are constructed and stored in the long-term memory. After the failure control mechanism shift search to another problems space, and new control structures for this process are created in the short term memory. With additional features of the problem generated in previous runs the new search has greater chances to succeed. However, this explanation may be applied to typical attempts of solving a problem by using several different strategies, without any Aha! experience. Only recently neuroscience has provided a deeper understanding of the insight phenomenon.

Studies using functional MRI and EEG techniques contrasted insight with analytical problem solving that did not required insight [27]. 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

Insight, a sudden Aha! understanding experience, occurs after an impasse, a period of frustration, and is preceded by feeling of imminence of solution and strong emotions, before the solution becomes clear. Strong involvement of right hemisphere of the brain has been observed during problem solving when insight occurred.

higher-level abstractions that can facilitate indirect associations. About 300 ms before insights occurred bursts of gamma activity has been observed. This has been interpreted by the authors as “making connections across distantly related information during comprehension ... that allow them to see connections that previously eluded them” ([27], p. 326). Bowden et. al [28] performed a series of fMRI experiments, confirming these results. In this interpretation initial impasse is due to the inability of left hemisphere, focused on the problem, to make progress. This deadlock is removed when the less-focused right hemisphere adds relevant information, allowing new associations to be formed. Aha! experience may result from activation by the pre-existing weak solution in the right hemisphere, suddenly reaching consciousness when the activation of the left hemisphere is decreased. Although these observations are important their explanation is rather nebulous. To understand the insight phenomenon first representation of words, concepts and the whole problem statement in the brain should be elucidated.

Words in the brain are an abstraction of acoustic speech input, changed into phonological, categorical representation. Categorical auditory perception enables understanding of a speaker-independent speech and is more reliable in a noisy environment. Phonemes, quantized building blocks of phonological representations (typically about 30-50 in most languages) are linked together in ordered strings by resonant states that represent word forms. In brains of people who can read and write, strictly unimodal visual representation of words in the Visual Word Form Area (VWFA) in the left occipitotemporal sulcus has been found [29]. Adjacent lateral inferotemporal multimodal area (LIMA) reacts to both auditory and visual stimulation and has cross-modal phonemic and lexical links. It is quite likely that the homolog of the VWFA in the auditory stream is located in the left anterior superior temporal sulcus; this area shows reduced activity in developmental dyslexics. In the Broca’s area in the frontal lobe precise motor representations that generate speech are stored. All these represen-

tations of word forms help to focus thinking processes. Activations of word forms are correlated with activity of other brain circuits, pointing to some experiences, perceptions and actions that define the meaning of words. Polysemic words probably have a single phonological representation and differ only by semantic extension. Analysis of the N200 feature of auditory event-related potentials shows that phonological processing precedes semantic activations by about 90 ms [30]. Similar phonological word forms activate adjacent resonant microcircuits. To recognize a word in a conscious way activity of its subnetwork must win a competition for an access to the working memory [31]-[34]. Hearing a word activates strings of phonemes, priming (increasing the activity) all candidate words and non-word combinations. 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 network. Semantic and phonological similarities between words should lead to similar patterns of brain activations for these words.

Language is lateralized usually in left hemisphere (LH), with right hemisphere (RH) responsible for largely non-verbal processing of speech information and recognition of limited number of words [31]. Right hemisphere is strongly connected to the LH, but such long projections cannot carry precise information about activations in the word form and extended word representation areas. RH may thus generalize over similar “semantic field” activations forming concepts at a high level of abstraction. Although these concepts have no names, as they are not associated with any word-form activation, they are very helpful in making inferences necessary to understand language. Simple inferences may be done locally through associative mechanisms in the LH, but more elaborate inferences relay on RH activations, involving especially the right temporal gyrus [35]. This conjecture is confirmed by a large psycholinguistic literature on the patients with RH damages, and similar conclusions from functional imaging of normal people: “LH may focally activate the semantic network, while RH activation may be more diffuse, coactivating more distantly related concepts” [36]. Distributed activations in the RH form various configurations that should activate back some regions in the left hemisphere, enabling to capture complex relations inherent in large semantic fields for concepts that have no name but are useful in reasoning and understanding. For example, “left eye” sounds correct, but “left liver” sound strange. The feeling of understanding is a kind of readiness potential of the brain to signal that inference processes due to the interplay of the left-right hemispheres have successfully finished. Associations at higher level of abstraction in the RH are passed back to facilitate LH activations that form intermediate steps in language interpretation. High-activity gamma burst project to the left hemisphere will prime subnetworks with sufficient strength to form associative connections linking the problem statement with partial or final solution. This is a universal mechanism that operates in case of difficult problems as well as in understanding of complex sentences.

High-activity gamma bursts, observed in the insight experiments [28], influence the left hemisphere priming larger subnetworks with sufficient strength to form associative connections that link problem statement through a series of intermediate transitions to partial or final solution. Such solutions may initially be difficult to justify, therefore the feeling of vague but imminent understanding is generated, replaced by real understanding when all intermediate steps are correctly linked. The solution may be surprising, based on quite different idea than initially entertained. Gamma bursts also activate emotions increasing plasticity of the cortex and facilitating formation of new associations. Emotional reaction should be proportional to the difficulty of forming new associations, therefore grasping a new difficult concept in a discussion generates only a mild reaction, while solving a difficult problem generates strong emotions, activating the reward system.

What computational inspirations may be drawn from these observations? One approach to model insight processes is based on small-world network analysis at the graph-theoretic level [37]. Activation of the RH during insight may create shortcuts between different subnetworks with dense local connections (small-

world subnetworks). The qualitative picture is quite clear: words and their associations correspond to patterns of activations that activate more general concepts in hierarchical way and part of the processing proceeds at non-verbal, high level of abstraction. The main challenge is how to use inspirations from neurocognitive linguistics to create practical algorithms for Natural Language Processing (NLP) and problem solving. It may be necessary to forget the details and look at the high-level, non-conceptual description of the problem. This process has distant analogy to reasoning at higher level of ontology and resembles the process of abstraction, that is formulation of general concepts by rejecting inessential details, very common in mathematics. Disambiguation and understanding of concepts requires extensive *a priori* knowledge that should be gained preferably from textbooks and structured knowledge sources. This reference knowledge may be modeled in several ways. The spreading activation

Understanding of words requires not only spreading activation through associations but also larger “semantic receptive fields” that activate neurons in the right hemisphere of the brain and usually do not have linguistic labels. Reasoning also involves concepts at different level of abstraction.

networks [38] could in principle provide most faithful models, but realistic large scale networks of this sort have so far not been created. These networks include both excitation and inhibition in the spreading activation process and are generalization of semantic networks [39]. Linguistic concepts should approximate word-form and semantic field activations in the brain, therefore connectionist models should not use nodes that represent whole concepts, but rather a fine-grained information about construction of words, such as morphemes or syllables. Context analysis will then provide guidance for spreading activation. Clusterization or granular computing techniques may try to capture similarities between semantic field activations and create hidden, internal concepts that correspond to the right-hemisphere activity, helping to make inferences during text comprehension.

Language development is grounded in internal representations of objects formed by the brain using information derived from perception, creating non-trivial semantic fields. To what extent this process may be approximated without embodied cognition? The main difficulty with neurocognitive approach to NLP is the lack of structural descriptions of common objects and concepts. Even the simplest concepts, such as those related to animals, do not have good description in the dictionaries, making creation of semantic memories from machine-readable sources quite difficult [40]. For example, everyone knows how a horse looks like, but a dictionary definition “solid-hoofed herbivorous quadruped domesticated since prehistoric times” (Wordnet), is certainly not sufficient to create correct associations. There are many proposals how to gain the missing knowledge from ontologies, dictionaries, encyclopedias, collaborative projects (MindNet, ConceptNet, Open Mind Common Sense Project), active search for possible relations between different concepts [40] and active dialogues in word games to add missing knowledge [41]. Statistical NLP approaches are based on the vector model, with different normalization methods that change word frequencies into useful features [42]. Vector approach may be treated as a snapshot of the network activity after several steps of spreading activation. In document categorization *a priori* knowledge may then be stored in reference vectors derived from description of concepts, for example names of disease that documents are related to [43]. Fuzzy prototypes may also be used instead of single reference vectors. To simulate spreading activation semantic smoothing techniques may be used, adding activation to concepts that are related to those discovered in the text. Synonyms and their parent concepts that are higher in ontological hierarchy should be added in the first place. In effect documents that may use quite different words are clustered into correct topics [44]. Knowing the topic and semantic types of concepts helps to disambiguate the meaning and annotate the text correctly. The activation of the semantic fields in the whole network has to be consistent, leading to the idea of active subnetworks, represented by graphs of consistent concepts [45]. These graphs should capture relations between concepts in their specific meanings, inhibiting alternative interpretations.

The lack of structural descriptions of common objects and concepts in lexical resources is a major problem in natural language processing. Statistical techniques do not solve this problem. Many experts place their hopes for progress in this area in embodied cognition. In principle this could create proper internal representations but development of good semantic memory may still be an easier way to understand language.

Brains are the only known systems capable of understanding natural language. Brain-like representations of linguistic concepts at morphological level have unique properties that facilitate various inferences needed to understand text. Although linguists are aware of the importance of the neurocognitive basis of language so far their interest has been restricted only to description of specific linguistic phenomena [46]. We do understand what has been missing in logical and statistical NLP approaches. This is a very fertile area for computational intelligence, with a lot of effort needed to create useful large-scale practical algorithms to approximate dynamical processes involved in language comprehension and production and to create good semantic memories. This may still be faster and easier way towards linguistic competence than embodied cognition.

Imagination and Creativity

Creativity is one of the most mysterious aspects of the human mind. Research on creativity has been pursued by educators, psychologists and philosophers. MIT Encyclopedia of Cognitive Sciences [8], Encyclopedia of Creativity [47] and Handbook of Human Creativity [48] described stages of creative problem solving and tests that can be used to assess creativity, but do not mention brain mechanisms or computational models of creative processes. Sternberg [48] has defined creativity 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. Brain processes behind creative thinking should not be much different from processes responsible for intuition and insight [49]-[51].

High intelligence is not sufficient for creativity although it is quite likely that both have similar neurobiological basis. Relationships between creativity and associative memory processes have been discussed already in [52]. Rich network of associations as well as strong right hemisphere involvement are clearly a pre-requisite for creativity. Heilman *et. al* agree that “creative innovation might require the co-activation and communication between regions of the brain that ordinarily are not strongly connected”, binding “different forms of knowledge, stored in separate cortical modules that have not been previously associated” ([53], p. 369). However, these authors do not consider lateralization of brain functions. One of the most important techniques in experimental psychology is aimed at investigation of priming effects [54]. The pair-wise word association techniques [54] measure response times to presentation of verbal stimuli, and are the most direct way to analyze associations among local networks coding different concepts. Priming techniques are based on cues that have influence on responses. Associations may differ depending on the type of priming (semantic or phonological cues), structure of the brain network that codes concepts, the activity arousal due to priming, and many other factors. Creative people should show 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 in this case show longer latency times, proportional to the difficulty of making an association (presumably related to the probability of making the transition between the two concepts, a good measure of distance in neural space).

Creativity is “the capacity to create a solution that is both novel and appropriate”. It is manifested in everyday activities, including language comprehension and production.

In one priming experiment [55] people with high and low scores in creativity tests saw the first word, followed for a brief (200 ms) moment by the priming cue (word), before the second word of the pair was displayed. More creative people indeed show greater ability to notice associations, especially for more difficult associations that less creative people frequently fail to notice. It should be expected that in a network that has small-world structure higher creativity means that there are more connections and thus higher transition probabilities between different subnetworks are possible. However, priming effects should differ, depending on the type of the priming cue. For easy associations positive priming (words that have related meaning) should lead to faster associations in all cases. Neutral priming, based on nonsensical or unrelated words, may in this case create in densely connected network (creative people) spread of activation in too many directions and thus competition for access to the working memory that will slow down the response times. In networks with fewer connections (less creative people) activation will mostly be spread through connections that correspond to easy associations, making the responses faster. When associations become difficult indirect activation routes are needed to facilitate transitions, perhaps involving inter-hemispheric transfers of activations. They are too weak or non-existing in less creative people and thus priming will not help them. Weak connections that exist in more creative brains may not be sufficient to facilitate quick transitions between the two paired word representations, but adding neural noise via nonsensical priming may increase the chance of such transitions. This is an example of the stochastic resonance phenomenon [56] that has been reported in visual, auditory and tactile perception, but evidently can also be noticed in associative thinking. Adding positive priming based on spelling activates only phonological representations close to that of the second word, therefore the influence should be weaker. All these effects have indeed been observed [55]. Experiments did not analyze the overlap between nonsensical words and the pair of words given for association at phonological and grapheme level, although this may reveal the microstructure of the associative process. These results support the idea that creativity relies on the associative memory, and in particular on the ability to link together distant concepts.

The first ingredient needed for creativity is thus sufficiently rich associative network, neural space capable of supporting complex states. High intelligence does not guarantee creativity. The second ingredient is imagination. Mental imagery is a well established field with its own *Journal of Mental Imagery*, started in 1977. Brains try to make sense of subtle cues, forming in parallel many hypothesis that compete with each other. Replacing a part of the spoken word by noise is sufficient to create an impression that the actual word that fits to the later context has actually been heard. For example [57], in the phrase “[noise]eel is on the —”, where the last word resolving the context is either “axle”, “shoe”, “orange” or “table”, incomplete information at the phonemic level is restored in the brain and the first word is heard as “wheel”, “heel”, “peel”, or “meal”, accordingly. The information that is consciously experienced is integrated in rather broad temporal window [30]. We are aware of the final result of the massive competition between various resonant states forming shorter and longer chains of activation, waiting for additional cues in form of context to resolve ambiguities. In absence of such cues some stronger activations temporarily win the competition, popping up in the working memory. It is quite likely that working memory is not a separate subsystem, but simply an

First thing needed for creativity is rich associative network, neural space capable of supporting complex states.

active part of the long-term memory (LTM) network due to priming and spreading of neural activation (see the review of the evidence for this point of view in [58]). 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 frontal lobe central executive areas (presumably this is the part we are conscious of) and some may be activated, but outside of this focus. Imagination depends on associations that the neural space is able to provide, but also on the energy, inner drive that may be due to the strong coupling via the dopamine projections between the frontal lobes and basal ganglia.

Second thing needed for creativity is imagination, the ability to combine together in many ways local brain activations in larger coherent wholes.

The final ingredient needed for creativity is the filtering system that selects most interesting (from emotional or cognitive point of view) mental images. Creativity is therefore 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 become highly active, reinforce mutually their activity, and form many transient configurations, inhibiting at the same time other activations. Difficult problems require long incubation periods that may be followed by an impasse and despair period, when inhibition lowers the 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. Only the most interesting associations (from the point of view of current goals) 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.

Third thing needed for creativity is filtering, granting access to the working memory of only the most interesting products of imagination.

Very few computational models addressing creativity have been proposed so far, the most interesting being Copycat, Metacat, and Magnificat developed in the lab of Hofstadter [59][60]. 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 this 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. Although this model has not been directly inspired by neurocognitive considerations it may approximate some fundamental processes of creative intelligence. The main application so far was in the design of new font families [60].

Results of experimental and theoretical research lead to the following conclusions:

- 1) 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:
- 2) distributed fluctuating (chaotic) neural activity, constrained by the strength of associations between subnetworks coding different words or concepts, responsible for imagination;
- 3) 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).

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 new words.

Neurocognitive approach to the use of words and symbols should draw inspirations from experimental psychology and brain research and help to understand putative brain processes responsible for creativity manifested in novel word creation. This could be a good area for more precise tests of creative processes using computational, theoretical and experimental approaches. Interesting names for products and companies are always in great demand. In languages with rich morphological and phonological compositionality (such as latin or slavic family of languages) novel words that cannot be found in the dictionary may appear in normal conversation (and more frequently in poetry). Although these words are newly invented their morphology gives sufficient information to make them understandable in most cases even without hearing the 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 desired 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 the more demanding IQ tests.

The simplest domain for testing models of creativity is creation of novel words, appropriate for web sites, names of the products or companies.

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 satisfied. 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 support 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 easier to construct using second and higher-order statistics for combination of phonemes (in some languages even combination of letters is acceptable, as spoken and written words are in close correspondence). 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.

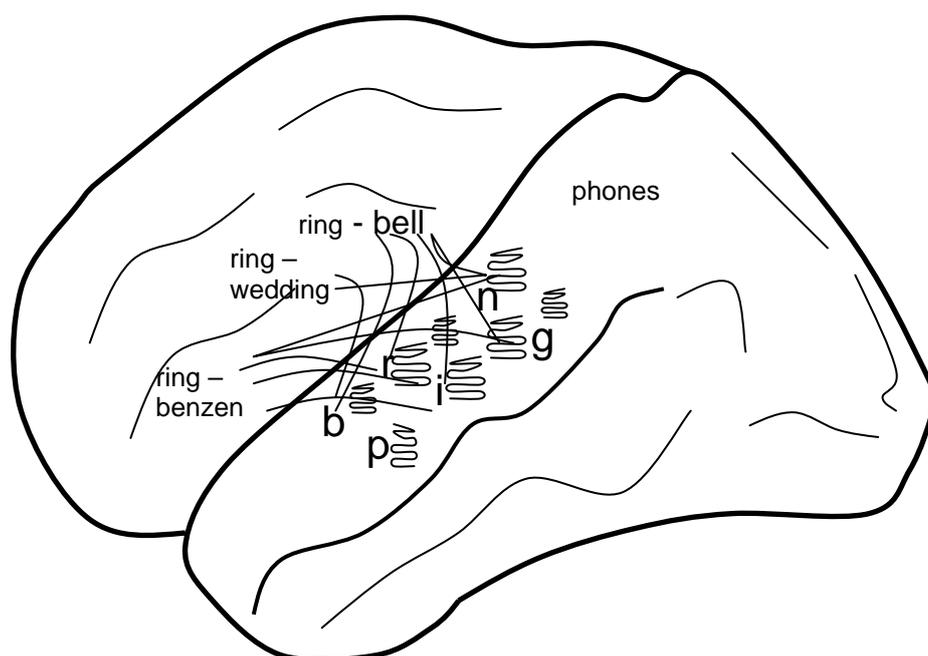


Fig. 3. Phonetic word-form „ring” has different extended representations.

Implementation of these ideas in a large scale neural model is possible, but as a first step simplest approximations have been tried [49]-[51]. The algorithm involves 3 major components:

- 1) an autoassociative memory (AM) structure, constructed for the whole lexicon of a given language at the morphological level to capture its statistical properties; it stores the background knowledge that is modified (primed) by keywords;
- 2) imagination implemented by forming new strings from combinations of substrings found in keywords (and their synonyms) 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 version a binary correlation matrix [61] has been used as autoassociative memory using single letters represented by temperature coding. Experiments with such 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 candidate words. There are several simple extensions to this model, either at the level of word representations, more complex network models, or learning algorithms. The first possibility has been explored to keep the algorithm as simple as possible. The list of elementary units has been 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 creation of sparse correlation matrix providing stronger language model constraints. Word is converted into a string of morphological atoms. To reflect constraints for filtering novel lexical strings binary weights may be replaced by correlation probabilities, taking into account word frequencies. The correlation matrix \mathbf{W} is calculated and normalized by dividing its elements W_{ij} by the sum of all elements in a row. Other ways to normalize this matrix have also been included in the program, or example additional position-dependent weights may stress the importance of the beginning and end atoms in words.

In the mental imagination step various combinations of atoms should be considered. As the number of combinations grows rapidly sequential filtering is used, combining pairs first and adding more atomic components to highly probable combinations only. 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 (wordnet.princeton.edu) synsets (sets of synonyms) 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 $\mathbf{W} + \lambda \mathbf{W}_p$, where \mathbf{W}_p is the weight matrix constructed only from the keywords. \mathbf{W}_p is multiplied by a factor λ that controls the strength of the priming effect. Using very large λ 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. 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. In current implementation of our algorithm such words are ranked low only in the final stage, when the “relevance” and “interestingness” filters are 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. 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

“Semantic density” around a novel word is estimated as the number of potential associations with words from the lexicon; this is taken as a measure of how interesting the new word appears to be.

lexical tokens or morphemes. For longer morphemes general similarity to other morphemes (rather than string equivalence) is 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 should be regarded 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. Knowing individual preferences and favorite expressions the algorithm could be to some degree personalized.

A few examples of results from such algorithm are presented here. First, interesting names for a website offering *shoes* are searched for. From the company brochure the priming set of keywords is extracted, consisting of such words as “running, sport, youth, health, freedom, air”. Several variants of the extended ngram model produced the following words: *shoebie*, *airenet*, *runnyme*, *sportip*, *windway*, *funkine*, *runnyme*, *moveman*, *runably*, *sporist*, *runniess*. 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. 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 summarizes the results, quoting approximate number of entries in Google search engine at the end of the year 2006.

Table I: Summary of interesting words related to *shoes*.

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

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 top words discovered in this case included *ideates*, *smartne*, *inveney*, *timepie*, *taleny*, *crealin*, *invelin*, *visionet*. Starting from an extended list of keywords: “portal, imagination, creativity, journey, discovery, travel, time, space, infinite”, more interesting words have been generated, with about $\frac{3}{4}$ already used as company or domain names. For example, *creativ* is used by creativ.com, *creativity* is used by creativity.com. Some words have been used only a few times (according to the Google search engine), for example *discovery* that can be derived from: disc, disco, discover, verity, discovery, creativity, verity, and may mean discovery of something true (verity). Another interesting word found 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. This example shows the importance of using extended keywords. Unfortunately novel words in the Internet get immediate attention of companies that try to reserve them for web sites.

In the near future we plan to create a web server for creation of novel words starting from short descriptions.

Perspectives

In everyday life intuition, insight and creativity are used more often than logic. For many years AI efforts to understand higher cognitive functions have been dominated by logics, the whole 5th generation computer project was focused on logic, but the results were less than encouraging. Evidently this was like barking at the wrong tree. Neuro-cognitive approach to understanding of intuition, insight and creativity is build on the common set of ideas, and is capable to explain, at least in a qualitative way, many high cognition phenomena. Obviously it is still quite speculative and the actual implementations are still rather simplistic but it seems to open the door to modeling of creative thinking at least in the narrow

In everyday life intuition, insight and creativity are used more often than logic. AI focused on logic has been barking at the wrong tree.

domain of word creation. Brain imaging and electrophysiological studies of the 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 in the brain [31]-[35],[62][63]. Probing associations and transition probabilities between brain states using priming techniques [54],[55] should lead to better understanding what kind of associations are most relevant. Research program on creativity, insight and intuition 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.

Intuition is not difficult to explain, both in recognition and reasoning for solving problems. Understanding insight leads to interesting inspirations for natural language processing. Creativity requires prior knowledge of the domain, imagination, and filtering of interesting 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 subjective preferences evaluating similarity to what has been assessed as interesting. Starting from a series of portraits and working in a space that decomposes visual inputs to shape, color and movement primitives, in analogous way as the linguistic input is decomposed into morphological parts, it should be possible to come up with interesting novel variants of paintings. 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.

Creativity requires prior knowledge of the domain, imagination, and filtering of interesting results.

To estimate practical usefulness of algorithms based on these principles their results should be compared with human inventiveness in a larger number of cases. Humans can obvious evaluate results in a better way then 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 may be implemented at a different level of neurobiological approximations, from detailed neural models to simple statistical approaches. However, 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. With sufficiently rich concept representation natural language processing may progress quite far, alleviating the need to use embodiment for creation of internal linguistic representation.

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 [64] 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 [65]. Linking mind and brain at the level of neural models is a great challenge. This approach should include creative and intuitive processes. In any case neurocognitive informatics, taking inspirations from the high-level brain functions, should have at least as bright future as artificial neural networks, taking their inspirations from the lowest, single neuron level.

The last sentence from the paper by Turing [1] is still the best summary of our current situation: “We can only see a short distance ahead, but we can see plenty there that needs to be done.”

Acknowledgment: Support by the Polish Committee for Scientific Research, research grant 2005-2007, is gratefully acknowledged. Algorithms to create novel words were implemented and run by Maciej Pilichowski.

References

- [1] A. Turing, “Computing Machinery and Intelligence”. *Mind* vol. 49, pp. 433-460, 1950.
- [2] *IEEE computational Intelligence Magazine*, vol. 1(3), 2006.
- [3] J.R. Anderson, “Learning and memory”. Wiley, New York, 2nd Ed, 2000.
- [4] T. Serre, A. Oliva, and T. Poggio. “A Feedforward Architecture Accounts for Rapid Categorization”. *Proceedings of the National Academy of Sciences (PNAS)*, vol. 104(15), pp. 6424-6429, 2007.

- [5] C. Koch, "The Quest for Consciousness: A Neurobiological Approach". Roberts & Company Publishers, Greenwood Village, Co, 2004.
- [6] W. Duch, "Brain-inspired conscious computing architecture". *Journal of Mind and Behavior*, vol. 26(1-2), pp. 1-22, 2005.
- [7] B. Heinrich and T. Bugnyar, "Testing Problem Solving in Ravens: String-Pulling to Reach Food". *Ethology*, vol. 111(10), pp. 962-976, 2005.
- [8] R. Wilson and F. Keil, Eds. "MIT Encyclopedia of Cognitive Sciences". MIT Press, 1999.
- [9] M.D. Lieberman, "Intuition: A Social Cognitive Neuroscience Approach". *Psychological Bulletin*, vol. 126(1), pp. 109-137, 2000.
- [10] D.G. Myers, "Intuition: Its Powers and Perils." Yale University Press, New Haven, CT, 2002.
- [11] W. Duch, R. Setiono and J. Zurada, "Computational intelligence methods for understanding of data." *Proceedings of the IEEE* vol. 92(5), pp. 771-805, 2004.
- [12] W. Duch and M. Blachnik, "Fuzzy rule-based systems derived from similarity to prototypes." *Lecture Notes in Computer Science*, vol. 3316, pp. 912-917, 2004.
- [13] W. Duch, "Uncertainty of data, fuzzy membership functions, and multi-layer perceptrons." *IEEE Transactions on Neural Networks*, vol. 16(1), pp. 10-23, 2005.
- [14] W. Duch, "Rules, Similarity, and Threshold Logic." *Commentary on E.M. Pothos, The Rules versus Similarity distinction. Behavioral and Brain Sciences*, vol. 28(1), pp. 23-23, 2005.
- [15] K. Grąbczewski and W. Duch, "Heterogenous forests of decision trees." *Lecture Notes in Computer Science* vol. 2415, pp. 504-509, 2002.
- [16] H.A. Simon, "Explaining the ineffable: AI on the topics of intuition, insight and inspiration." *Proc. 14th Int. Joint Conference on Artificial Intelligence*, vol. 1, pp. 939-948, 1995.
- [17] A. Linhares, "An Active Symbols Theory of Chess Intuition". *Minds and Machines*, vol. 15(2), pp. 131-181, 2005.
- [18] A. Newel, "Unified theories of cognition." Harvard Univ. Press, Cambridge, MA, 1990.
- [19] J. Hawkins and D. George, "Hierarchical Temporal Memory – Concepts, Theory, and Terminology." *Numenta Inc.*, www.numenta.com, 2006.
- [20] O. G. Selfridge and U. Neisser, "Pattern recognition by machine." *Scientific American*, vol. 203, pp. 60-67, 1960.
- [21] D.E. Rumelhart and J.L. McClelland (eds), "Parallel Distributed Processing, Vol. 1: Foundations, MIT Press, Cambridge, MA, 1986.
- [22] G.E. Hinton, S. Osindero and Y. The, "A fast learning algorithm for deep belief nets." *Neural Computation*, vol. 18, pp. 381-414, 2006.
- [23] W. Duch and G.H.F. Diercksen, "Feature Space Mapping as a universal adaptive system." *Computer Physics Communications* vol. 87, pp. 341-371, 1995.
- [24] R.J. Sternberg, and J.E. Davidson, "The nature of insight". MIT Press, Cambridge, MA, 1995.
- [25] E.A. Feigenbaum and H.A. Simon, "EPAM-like models of recognition and learning." *Cognitive Science*, vol. 8, pp. 305-336, 1984.
- [26] H.A. Simon, "The sciences of the artificial." 2nd ed, MIT Press, Cambridge, MA, 1981.
- [27] M. Jung-Beeman, E.M. Bowden, J. Haberman, J.L. Frymiare, S. Arambel-Liu, R. Greenblatt, P.J. Reber, and J. Kounios, "Neural activity when people solve verbal problems with insight." *PLoS Biology*, vol. 2(4), pp. 500-510, 2004.
- [28] E.M. Bowden, M. Jung-Beeman, J. Fleck, and J. Kounios, "New approaches to demystifying insight." *Trends in Cognitive Science*, vol. 9, 322-328, 2005.
- [29] S. Dehaene, L. Cohen, M. Sigman and F. Vinckier, "The neural code for written words: a proposal." *Trends in Cognitive Science*, vol. 9, pp. 335-341, 2005.
- [30] S. Grossberg, Resonant neural dynamics of speech perception. *Journal of Phonetics*, vol. 31, pp. 423-445, 2003.
- [31] F. Pulvermüller, *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge, UK: Cambridge University Press, 2003.
- [32] F. Pulvermüller, Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, vol. 5, pp. 517-524, 2001.
- [33] F. Pulvermüller, A brain perspective on language mechanisms: from discrete neuronal ensembles to serial order. *Progress in Neurobiology* vol. 67, pp. 85-111, 2002.
- [34] 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.
- [35] S. Virtue, J. Haberman, Z. Clancy, T. Parrish and M. Jung-Beeman, "Neural activity of inferences during story comprehension." *Brain Research*, vol. 1084, pp. 104-114, 2006.

- [36] K.I. Taylor Monsch, "Semantic Language in the Right Hemisphere: Divided Visual Field and Functional Imaging Studies of Reading." PhD Thesis, University of Zurich, 2002.
- [37] M.A. Schilling, "A 'Small-World' Network Model of Cognitive Insight." *Creativity Research Journal*, vol. 17(2-3), pp. 131–154, 2005.
- [38] F. Crestani, Application of Spreading Activation Techniques in In-formation Retrieval. *Artificial Intelligence Review* 11:453-482, 1997.
- [39] J.F. Sowa, ed. "Principles of Semantic Networks: Explorations in the Representation of Knowledge." Morgan Kaufmann Publishers, San Mateo, CA, 1991.
- [40] J. Szymanski, T. Sarnatowicz and W. Duch, "Towards Avatars with Artificial Minds: Role of Semantic Memory." *Journal of Ubiquitous Computing and Intelligence*, American Scientific Publishers (in print).
- [41] J. Szymanski and W. Duch, Semantic Memory Knowledge Acquisition Through Active Dialogues. Proc. of the 20th Int. Joint Conference on Neural Networks (IJCNN), Orlando, IEEE Press, August 2007 (in print).
- [42] C.D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing* MIT Press, Cambridge, MA 1999.
- [43] L. Itert, W. Duch and J. Pestian, "Influence of a priori Knowledge on Medical Document Categorization." IEEE Symposium on Computational Intelligence in Data Mining, IEEE Press, April 2007, pp. xxx.
- [44] W. Duch, P. Matykiewicz and J. Pestian, "Neurolinguistic Approach to Vector Representation of Medical Concepts." Proc. of the 20th Int. Joint Conference on Neural Networks (IJCNN), Orlando, IEEE Press, August 2007 (in print).
- [45] P. Matykiewicz, W. Duch and J. Pestian, "Nonambiguous Concept Mapping in Medical Domain." *Lecture Notes in Artificial Intelligence* vol. 4029, pp. 941-950, 2006.
- [46] S. Lamb, *Pathways of the Brain: The Neurocognitive Basis of Language*. Amsterdam & Philadelphia: J. Benjamins Publishing Co. 1999.
- [47] M. Runco and S. Pritzke, Eds. *Encyclopedia of creativity*, vol. 1-2, Elsevier, 2005.
- [48] R.J. Sternberg, Ed. *Handbook of Human Creativity*. Cambridge: Cambridge University Press, 1998.
- [49] W. Duch, "Computational creativity." IEEE World Congress on Computational Intelligence, Vancouver, July, 16-21, IEEE Press, pp. 1162-1169, 2006.
- [50] W. Duch, "Creativity and the Brain." In: *A Handbook of Creativity for Teachers*. Ed. Ai-Girl Tan, World Scientific Publishing 2007 (in print).
- [51] W. Duch and M. Pilichowski, Experiments with computational creativity. *Neural Information Processing – Letters and Reviews*, Vol. 11, No. 3, March 2007 (in print).
- [52] S.A. Mednick, "The associative basis of the creative process." *Psychological Review*, vol. 69, pp. 220–232, 1962.
- [53] K.M. Heilman, S.E. Nadeau and D.O. Beversdorf, "Creative innovation: Possible brain mechanism." *Neurocase*, vol. 9, pp. 369-379, 2003.
- [54] T.P. McNamara, *Semantic Priming. Perspectives from Memory and Word Recognition*. Psychology Press, New York, 2005.
- [55] A. Gruszka and E. Necka, Priming and Acceptance of Close and Remote Associations by Creative and Less Creative People. *Creativity Research Journal* Vol. 14(2), pp. 193-205, 2002.
- [56] T. Wellens, V. Shatokhin, and A. Buchleitner, Stochastic resonance. *Reports on Progress in Physics* Vol. 67, pp. 45-105, 2004.
- [57] R.M. Warren and R.P. Warren, "Auditory illusions and confusions." *Scientific American*, vol. 223, pp. 30–36, 1970.
- [58] 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.
- [59] D.R. Hofstadter, *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. NY: Basic Books, 1995.
- [60] J. Rehling, *Letter Spirit (Part Two): Modeling Creativity in a Visual Domain*. PhD Thesis, Indiana University, 2001.
- [61] T. Kohonen, Correlation matrix memories, *IEEE Transactions on Computers C*, Vol. 21, pp. 353-359, 1972.
- [62] 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.
- [63] A. Martin, C.L. Wiggs, L.G. Ungerleider and J.V. Haxby, Neural correlates of category-specific knowledge. *Nature* 379, 649-652, 1996.

- [64] 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.
- [65] 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.

Wlodzislaw 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)