Preconceptual Creativity

Tapio Takala
Department of Computer Science
Aalto University, Finland
tapio.takala@aalto.fi

Abstract

Creativity, whether seen in personal or historical scope, is always relative, subject to the contextual expectations of an observer. From the point of view of a creative agent, such expectations can be seen as soft constraints that must be violated in order to be deemed as creative. In the present work, learned conventions are modeled as emergent activity clusters (pre-concepts) in a self-organizing memory. That is used as a framework to model such phenomena as stereotypical categorization and mental inertia which restrain the mind when searching for new solutions. Using the kinematics of a robotic hand as an example, the models' dynamic behavior demonstrates primitive creativity without symbolic reasoning. The model suggests cognitive mechanisms that potentially explain how expectations are formed and under which conditions an agent is able to break out of them and surprise itself.

Creativity is in the Eye of Beholder

Creativity is a concept that defies exact definition. The commonly accepted view that creativity is a process resulting in novel and useful products (Mumford 2003) appears to be loose, because in the strict sense even a slightest modification would make the product novel. Another often cited definition is by Newell et al. (1959) who generously view it as a problem-solving process presenting one or more of the following: novelty and value, unconventional thinking, high motivation, and ill-defined problems. They continue by admitting that no more specific criteria can be set for separating creative from non-creative thought processes.

Surprise, more or less as a synonym of unconventional or unexpected, is often considered a necessary condition for creativity (e.g. Boden 1990). However, it may be difficult to distinguish unconventional from mere novelty, as it depends on the observers' subjective experience and conventions. Moreover, novelty is a moving target: once an invention is made it becomes legacy – unless it is forgotten and may be reinvented. Like Grace and Maher (2014), we conclude that creativity is in the eye of beholder, and cannot be defined objectively.

To get a grasp of the relative nature of creativity we adapt the generate-and-verify model by Newell et al. (1959) into variable scopes (Fig.1). The products of a generator (G) passing the evaluation (E) on one level are used as input to evaluation on the next level. A person using computer as a generator (G_p) may find designs passing her evaluation criteria (E_p), but while showing these to others she (together with her computer) acts as a generator (G_h) for the society where others collectively act as evaluators (E_h).

On the societal level creativity appears to be a statistical concept formed by opinions of the population under study. Czickszenmihaly (1997) studied individuals (G_h) with a reputation of being creative. Maher et al. (2013) studied the evaluation (E_h) with a temporal regression model of car designs, where outliers have higher potential for surprise and creativity.

In this paper we concentrate on the personal level (P-creativity), trying to computationally model some of the phenomena happening in a person's mind when a creative moment is encountered. In this respect the generative process is not in our focus. Although various control strategies (analogy, negation, metaphors, etc.) can make it more efficient and interesting, it may as well be a black box. Essential for creativity is the evaluation process, which recognizes value and novelty in products of the generator. It becomes surprised if something unexpected is produced, i.e. if its expectations are violated.

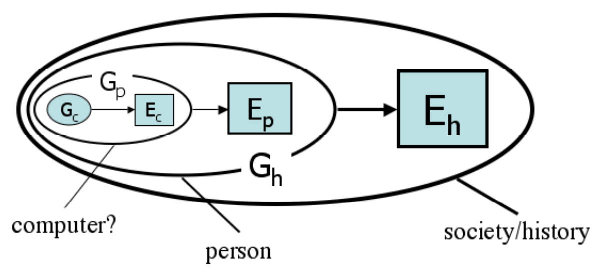


Figure 1. Context defines the expectations (E) against which the creativity of a generative process (G) is evaluated (from Takala, 2005).

What are the expectations then? They can be understood as constraints on the product (or process): what it should or is assumed to be (or how it is assumed to be done). They may be hard (defining the domain), such as laws of nature and logic or explicit rules of a game, but they can also be soft (acquired) constraints: habits, conventions, manners, fashion, social norms, political correctness, etc. These soft constraints are contextual and subject to consideration, applying in one situation but irrelevant in another. But they can be very hard in practice if based on psychological repression. This may serve as an interpretation of Boden's expression that creativity produces "previously impossible ideas". An *idée fixe*, or design fixation (Jansson and Smith 1991) may be the most common obstacle hindering creativity. Such soft constraints form the "box", out of which we are supposed to take a leap.

What makes creativity valuable is that it is a constructive, sense-making act, not just anarchy that randomly defies any rules without a purpose. The new act must in some (novel) way be regular and repeatable. Creativity is search for a constructive and consistent solution assuming some constraints but neglecting or modifying others. By and large, *creativity is management of constraints* for finding a resolution of conflicts among them.

Different degrees of creativity can be identified according to the level of abstraction, or cognitive complexity: (1) Most trivial, though subjectively surprising, is the case when a solution is already known but happens not to be in the current scope of attention: "It just didn't come to my mind". (2) Some effort is required if the solution is not familiar as such but is potentially reachable by known methods or rules. Then essential is the selection of right starting points and methods to proceed with, while neglecting the obvious ones that may distract the process. An example of this is the need to backtrack in order to avoid an obstacle instead of stubbornly pushing straight towards a goal. (3) Yet a higher level comes if the solution is potentially reachable within the hard constraints, but requires constructive actions on the metalevel, i.e. new rules or methods. (4) Finally, even if the product is actually not realizable, we may still act creatively by imagination, neglecting the physical constraints.

The first two degrees, interpreting unexpectedness as changes in the scope of attention (relaxing soft constraints and that way releasing latent possibilities), are demonstrated below using a self-organizing memory as a model. The higher levels, requiring symbolic rules to be changed, are out of the scope of this paper. So is the sometimes required property that creativity should reflect itself, consciously recognizing that something novel and valuable has been formed.

On Representations

What can be done (consciously acted on) in problem solving, depends on its conceptual representation. This is an important research issue for cognitive science. The main bulk of AI research concentrates on the symbolic level, dealing with logic, language and inference rules. Another end is the subsymbolic sensory area, dealing with neural networks, associative memory and statistical inference. The well-known frame problem, or symbol grounding, calls for connections between the two. In the present work, we are not trying to fill the gap fully, but approach it from bottom up, demonstrating how primitive conceptual representations possibly form from the embodied information.

As the enaction theory (Stewart et al 2011, Rosch et al. 1992) assumes, regularities of the world are learned by receiving repeated stimuli and doing explorative actions. Conditioning and mimicking are two basic psychological principles facilitating this. Later, abstractions of experiences form as subsymbolic concepts. They facilitate more efficient behavior as perceptions are immediately categorized into known classes that may trigger preprogrammed reactions.

Such predefined reactions are of advantage in the world where things are quite predictable. A repeatedly adequate behavior gradually becomes the expected, a rule to be followed. Novel reactions are necessary only if the conditions change – as the proverb says: "necessity is the mother of invention". From evolutionary perspective, however, it may also be of advantage to try out novelties even without a reason, to become prepared for changes. Such tendency is called curiosity, or creative personality.

In neural networks, the sensory information is modeled statistically as conditional distributions and associations. Connecting this to the higher cognitive processes has long been a challenge. Gärdenfors (2000) suggests *conceptual spaces* as a potential bridge between sensory and symbolic levels, a theory of concept formation on supersensory but subsymbolic level. The idea is to describe objects with their properties that act as dimensions of a geometric (metric or topological) space. Individual objects are represented as points in this space, and their generalized conceptual representations as (convex) areas. Inspired by prototype theory (Rosch 1973) Gärdenfors suggests that natural categories may be represented as a Voronoi tessellation around central points representing stereotypical prototypes. This way the extensional (set of experienced samples) is converted into a more efficient intensional (set of constraints) representation.

In this paper, a somewhat similar framework is built, though not relying on a geometric feature space like Gärdenfors (2000) and Chella et al. (2014), but letting the neural cells of a self-organizing network to serve as representative samples of the sensory input. Concepts are not formed explicitly but just as (dynamic) clusters of similar cells. Thus we call it *preconceptual*, resembling the development stage of mind before actual conceptual thinking, in which sensorimotor activity predominates. Pylyshin (2001) uses the term in a compatible manner to describe situated vision, referring to objects that are identified but not defined by their properties. The idea also closely relates to 'proto-symbols' by Brooks and Stein (1994), who use the term for patterns of behavior that represent generalizations but appear rather as signals than formal symbols. Creativity is then demonstrated in primitive form, i.e. problem solving and conflict management using implicit concepts without symbols (Brooks 1991).

Implementation with Self-Organizing Map

The computational framework we use is based on the Self-Organizing Map (SOM) by Kohonen (2001). It is a widely used clustering device in pattern recognition and data analysis. As a biologically motivated neural network it is an interesting model for cognitive science. It has been suggested by Gärdenfors (2000) as a means of implementing conceptual spaces, though his approach is rather programmatic than an actual implementation.

The SOM is a neural network consisting of an array of cells connected to a vector of input values (Fig.2). The connection weights w_{ij} of a cell are initially random but are changed as follows: Given an input vector $\mathbf{x}$, the cell with best matching weight vector $\boldsymbol{w_j}$ is selected, and its weights are tuned towards the input values. A similar tuning is also done in its neighbor cells.

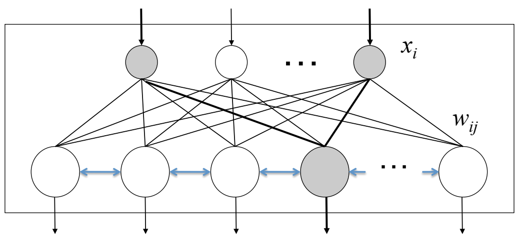


Fig. 2: The principle of SOM. Input vector $\boldsymbol{X}$ is compared with weight vectors $\boldsymbol{W_j}$ of the cells. The best matching unit is selected and its weight vector tuned towards the input in the training phase. As associative memory, SOM returns $\boldsymbol{W_j}$ as output in response to partial input (an example: active elements emphasized).

With a large number of input samples, the network organizes itself by unsupervised machine learning instead of using explicitly given concepts. Effectively it builds a model of the training input's statistical distribution, such that each cell represents a collection (a vector) of associated input values, and the number of cells with similar values reflects the density of those value combinations in the input. Usually SOM is implemented with low-dimensional topology (typically a regular 2-D array), and becomes folded if applied to higher dimensional input. An example is given in Figure 3.

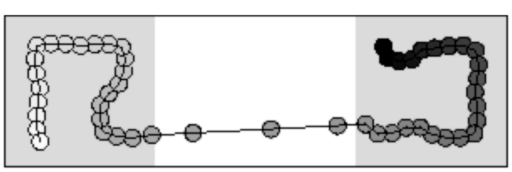


Fig. 3: One-dimensional SOM (chain of cells) trained with data from a 2-D distribution concentrated in the grey areas (cells are visualized in the input space in locations of their learned values).

In pattern recognition SOM is widely used as a classification device. It tells efficiently if a given input vector belongs to one category or another. This helps in data compression as complex input vectors can be quantized and represented with a smaller number of dimensions.

In SOM, similar cells emerge close to each other resulting in associations between a cell and its neighborhood. If there are concentrations in the input distribution, similar cells form clusters separated by dissimilar boundaries (Figs. 4a).

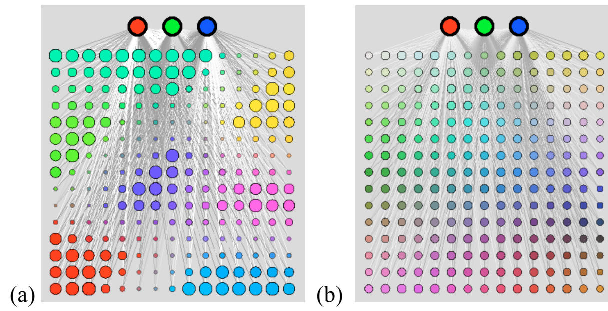


Fig. 4: A two-dimensional SOM trained with RGB values of (a) discrete colors (b) flat color spectrum. Cell color shows its learned values, cell size indicates similarity with its neighbors.

As each cell represents a vector of correlated input values, the SOM can act as an associative memory. A partial input (i.e. the values given for some inputs, and the rest undefined) as a stimulus activates the cells according to their similarity with the defined inputs. As result we get for each cell the probability of its value vector to become the output. Then we select the cell best matching the partial input, and take its weight vector as output (see Fig.2). Effectively the associative memory would fill in the undefined values by those from a cell selected by highest probability. Practical applications are found in image completion (Kohonen 2001), or information retrieval (Kohonen et al. 2000), for example.

The separable clusters (as in Fig. 4a) can be interpreted as primitive concept formation ("preconcepts"). When an input activates some cells, their similar neighbors are activated as well in the cluster. Then if the cluster were labeled with semantic information (such as color name), the input would be identified with that. The behavior resembles categorical perception in psychology (Goldstone and Hendrickson 2010) in the sense that the classification of any input within a cluster would get strong support by a group of cells, whereas an input falling to an area boundary would be in "unknown" territory where classification is unreliable. This coincides with the phenomenon in categorical perception that stimuli near category boundaries are more difficult to identify than within categories.

It is not clear if the human perceptual categories are independent of symbolic concepts, nor if they are presented by stereotypical prototypes or area boundaries. We hypothesize that it is possible to form concepts without higher level semantics, if such identifiable areas emerge. Such does not happen if the input distribution is flat without statistical foci (Fig. 4b).

A Case Study

In this section, we show how an associative SOM can be used to solve the control problem of a kinematic hand, and demonstrate preconceptual creative behavior in that context.

Setting the Scene

An articulated kinematic hand mechanism consists of a set of links connected at rotational joints to make a chain. In our case there are two such links (Fig. 5). Using the two joint angles (α and β) as motoric controls, the hand can reach points in the (*x, y*) plane within an area delimited by its physical constraints (i.e. the allowed ranges of control angles, and possible other geometric obstacles). The hand position can easily be calculated by trigonometry from the angles and lengths of the joined links, whereas the inverse is non-trivial. This inverse kinematics (IK) problem, finding control values for angles, given a target position, is generally a hard problem without analytical solution. A simple solution exists for our case with only two degrees of freedom, but it is still interesting due to its non-linearity (including singularities), physical constraints, and non-uniqueness of the solution: the same point can be reached by left or right handed configuration (negative or positive values of β, respectively).

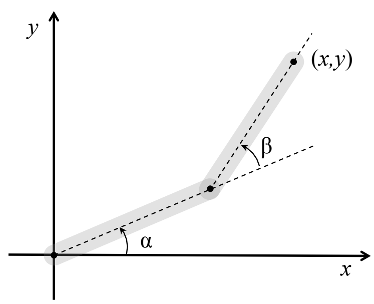
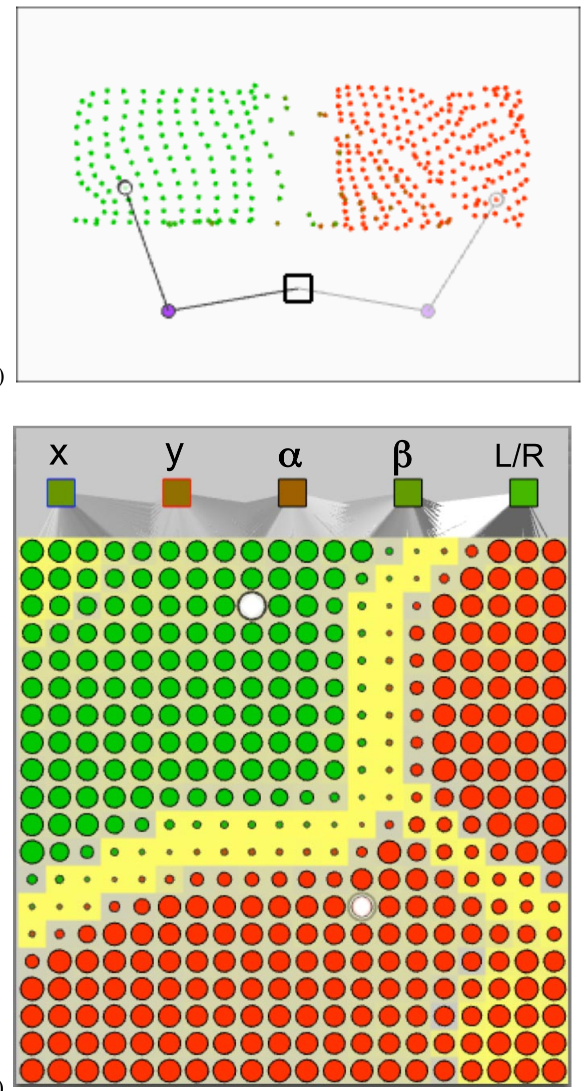


Fig. 5: Kinematics of a robotic hand

Among many other techniques, feed-forward neural networks have been proposed to solve the IK problem by training the system with random samples from the configuration space (e.g. Duka 2013). In case the problem is under-constrained (i.e. the robot has redundant degrees of freedom), sampling can be utilized to satisfy additional goals, such as moving in a certain style. Wiley and Hahn (1997) propose building from the given positions a resampled grid that serves as a geometric index, out of which the final angle-target combinations are calculated by interpolation. Our approach is similar to both of these in the sense that a neural network is trained to form a grid-like index, from which candidate starting points are selected for final approach to the target.

Let us assume our humanoid robot has two hands with their physical limits (hard constraints) similar to those of the human left and right hand. Each hand is trained to work in its most natural area (left/right in front of the base) as in Fig. 6a. The system is implemented in one SOM with two inputs for hand position (x, y), two for joint angles (α, β), and one (binary) input for handedness. Then clusters automatically form in SOM corresponding to left and right handed operation (Fig 6b). Their actual shape is random, sometimes bifurcated or consisting of multiple foci, but the areas are clearly identifiable. The clusters are separated by a boundary where the cells are less similar with their neighbors (shown in yellow).

(a)



(b)

Fig. 6: Training areas of hands (L=green, R=red) in the experiment. a) in robot space, b) as clusters formed in SOM. Two sample positions shown: white cells in SOM and the corresponding left (solid) and right (shadowed) hand positions.

The IK problem is solved by association, taking the target's coordinates as partial input, finding the cell(s) that best matches with it, and returning its weight values for the missing inputs (the control angles α and β):

$$f: (x, y, ?, ?) \longrightarrow (x', y', \alpha, \beta)$$

Although the result as such is not exact, it provides a good starting point for an iterative final approach. The movement direction needed in the iteration phase can be estimated from the cell's neighborhood by differentiation (approximating the Jacobian of parameters). This is a common strategy with actual robots and well grounded by biological action where proprioceptive memory and motor programs (Keele 1968) quickly lead to approximately right position and the final approach is done with the help of sensory feedback. In our implementation this phase is computed explicitly, but the Jacobian differentials could as well be learned by the SOM, if continuous movements instead of random positions were used in the training phase.

Targets within the trained areas are easily reached with the method above, and if the target is not too far out from the trained area, it usually can be reached from the closest starting point by the final iteration (Fig. 7a).

Acting Creatively

Now let us take a challenge where the simple approach does not work, by setting the target in a place not reachable from the closest point by direct iteration. This may be caused by a limitation of the mechanism itself or happen due to a physical obstacle (such as the box wall in Fig. 7c). Then the final approach gets stuck and we need to find a new starting point.

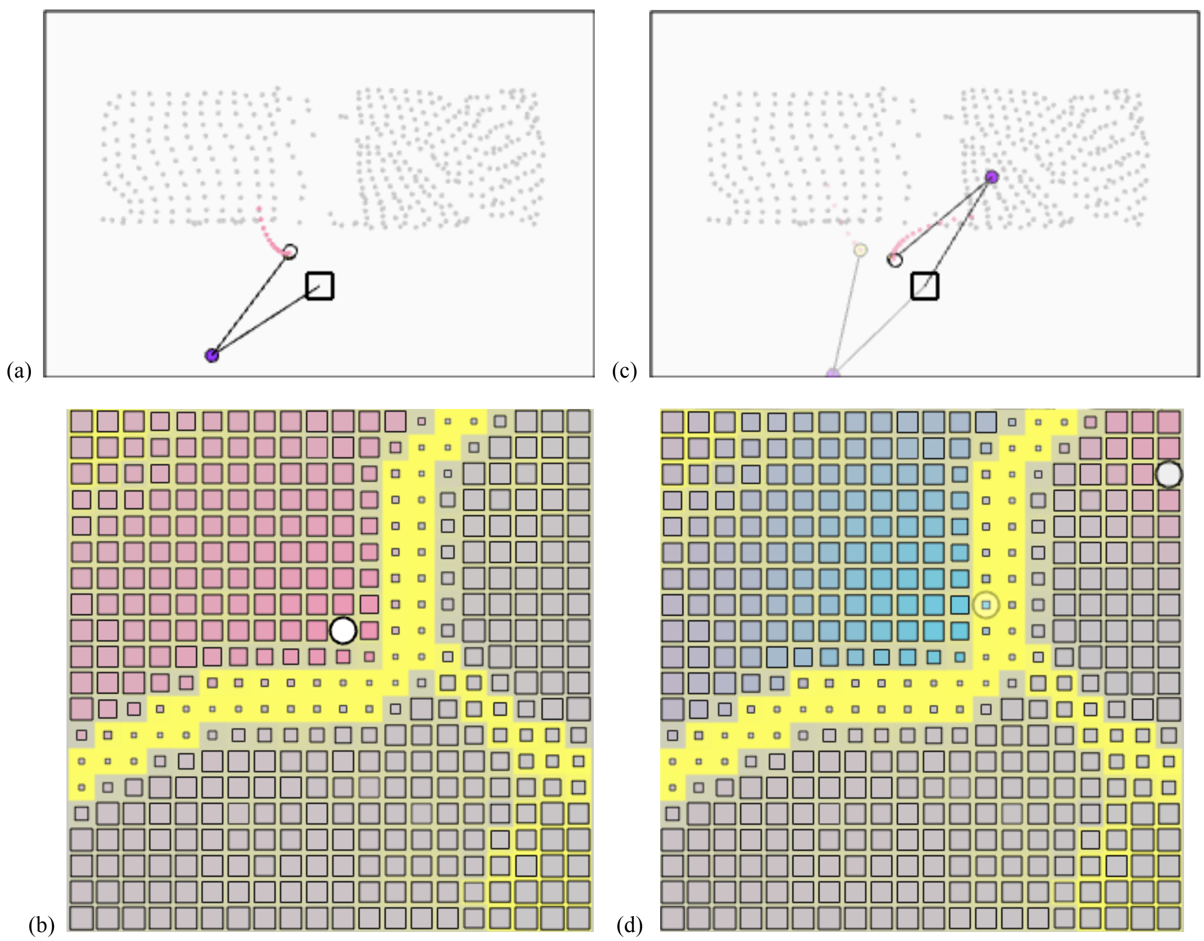


Fig. 7: Creativity in search for IK solutions. (a) target point reachable with left hand (final iterative approach shown as a sequence of red dots), (b) SOM cell (white circle), found in a recently active cluster (pink), defines the starting point for approach, (c) target appears impossible for the left hand due to the wall obstacle at its "elbow" (shadowed), but a new starting point feasible for the right hand is found (solid), (d) corresponding activity in SOM, where the previous starting point is surrounded by negative feedback effect (blue) due to unsuccessful trials, and the new point gets positive feedback (pink) which propagates to neighboring cells.

Though in principle any starting point could be considered as a new candidate, a random search is not very effective. Even if a cell's probability to be selected is weighted by its correlation with the input, a random method would mostly suggest candidates near the one which already lead to a dead end. The obvious engineering solution, trying out all candidate points in successive order, is not suitable here because sorting would call for higher level conceptual thinking and a different memory organization. We do not want to give the system any ready-made domain specific heuristics either, but want it to rely on very generic principles. As such an approach we utilize supervised reinforcement learning with a short-term memory (STM).

We implemented a distributed STM as an additional variable in each cell. It modulates the cell's probability to be selected as candidate for a trial. Its value would be increased by positive feedback from a successful case and decreased if the trial fails. Following the self-organization principles, these changes are also propagated to the cell's neighborhood but only among similar cells. To keep the operation dynamic, both positive and negative effects are gradually faded, possibly with different time constants.

The system's behavior now depends on its short-term history, its sensitivity to feedback, and the relative time constants. Let us assume the robot has operated for a while with targets in the left-hand area. Then the cells in the corresponding cluster(s) have been activated a lot, and due to positive feedback their probability to be selected again is high (pink color in Fig. 7a-b). When the target moves to a near but unreachable position (Fig 7c), the same cells continue to be activated as candidates, but a failure to reach the goal from one starting point will make the probability of that cell (and its close neighborhood) low. However, because of recent positive activity, the search will still continue with other cells in the same cluster. Then the further course of action is determined by the system's history and parameters as follows.

If a cluster's temporal activity is high (due to operating long in that area) and fading slower than the effects of negative feedback, the system will continue search within the same cluster despite of being unsuccessful. This corresponds to *mental inertia*, the tendency to keep on temporal preferences, i.e. the agent's expectation that a recently useful concept will continue to be so, an *idée fixe*.

However, if the negative feedback is more persistent and eventually dominates the whole cluster (indicated by blue color in Fig 7d), then a cell in some other cluster (probably one with next best correlation with the target) gets highest probability and will be taken as starting point for a trial. If it does not succeed, negative feedback will make its neighborhood less probable and the search continues somewhere else. Effectively this would implicitly perform an ordered search, though without explicit sorting.

Once a successful case is found (possibly requiring iterative final approach as in Fig. 7c), it will get positive feedback which is diffused to its neighbors in the same cluster, too (pink color in Fig. 7d). If the agent's operation continues with further targets nearby, this neighborhood will provide successful candidates again, and eventually the cluster becomes predominant: a primitive paradigm shift has happened, *heureka!*

Analysis of system behavior

We can evaluate the system theoretically and get the following qualitative observations, also confirmed by experiments with different parameters and test conditions.

In the above case, the creative leap was required because the left hand was unable to continue operation due to a constraint. Had the system a different history, with the right hand recently used before going to the new target, the new solution would have been obvious because of the predominant right hand: no creative moment, nothing unexpected, although new compared to what had been learned and stored in the long term memory (SOM). This is in alignment with the general observation that mental fluidity is induced by pressures (Hofstadter and Mitchell 1995) and may not happen otherwise.

Sticking with recently used behavior and building expectations is necessary for the system to act creatively, but it is not sufficient alone. Without negative feedback from an unsuccessful trial the system will keep trying the same over and over without getting anywhere.

Without any (positive or negative) feedback the system looses its temporal properties and reacts always the same way in a given situation, governed by the associative memory alone.

An interesting situation is encountered if we neglect the positive feedback but keep the negative. This leads to an "anti-sticking" behavior: once a cell has been used, neither it nor its close neighbors will be used for the next trial, but something loosely associated with the input. As the effect of negative feedback gradually fades away, the system may return to this cell if its association to the input is high, but only temporarily, and then jump to another cell. Overall, this resembles divergent thinking: variable alternatives are tried out, not randomly but guided by associations.

In our case study the robot's handedness was given as an explicit input feature to the SOM. This makes a clear distinction between clusters corresponding to left and right handed operation, respectively. However, this feature appears to be unnecessary, as similar behavior may emerge anyway if there only are two or more separate clusters formed from the distribution of input value combinations.

The ability to act creatively depends on the problem domain and its representation: if there are local optima where one may get stuck, there is a possibility for radical moves – otherwise a too simple route may lead to the solution. In this respect our system can be compared with optimization: Gradient search is a sticky strategy corresponding to the case with positive feedback only. Parallel search methods, such as genetic algorithms and simulated annealing, may lead to unexpected solutions, though in their basic form they have no such concept as surprise. However, the 'temperature' that makes simulated annealing process to look for more random options may well be compared to the negative feedback in our system.

Discussion

Different degrees of creativity, as mentioned in the introduction, can be demonstrated with our system. The case when a solution is already familiar (or reachable by iteration) but "didn't come to my mind" is modeled if the recent history has built strong temporal preference for a subset of solutions. This manifests itself as the agent's "sticky" tendency to sometimes utilize iterative approach from recently used starting points even if there were a better starting point stored in SOM, but this alternative is in a different cluster.

The more interesting case, target reachable within hard constraints but outside the most obvious trained area, is demonstrated when starting to use the other hand after trying and failing with one (as in Fig. 7). This can be interpreted as transformational creativity on preconceptual level, a change in the predominant cluster (rule) used in the agent's operation. It involves relaxation of soft constraints (giving up accustomed solutions), an essential property of creativity.

Whether this should be called creativity, may be an arguable question. Hristovski et al. (2011) have studied a similar situation of limb movements in the context of boxing. On the one hand, they state that any novel movement that has not been performed previously by an individual can be considered a P-creative act. On the other hand, they note that movement system bistability yields too much predictable behavior to account for creativity. Our case may be interpreted as the latter due to the binary choice of left or right hand in any situation, or the former because the exact hand movement is not predictable. A deeper analysis of the system's dynamics may be needed to take a stance.

Although our model shows qualitative changes in the robot's dynamical behavior, it is missing temporal anticipation, which could be utilized for creative planning of actions. The implementation as such does not support reasoning about an action's consequences that would be needed for goal-oriented behavior and higher-level expectations (Lorini and Falcone 2005). However, similar techniques might be used for learning temporal associations as well, thus making it a platform for further development.

Lorini and Falcone (2005) used formal logic to describe expectations and surprise in symbolic domain. At the other end of the scale, specific neural assemblies have been found that correspond to these phenomena in visual cognition (Egner et al. 2009). This suggests that a neural network model may be feasible. Gabora (2010) presents a schematized associative memory where neural cliques are alternatingly recruited for analytic and associative modes of thought, which is supposed to be essential for creativity. The model does not consider expectations and surprise, nor computational implementation, but the activation function of neurons may be comparable to our feedback mechanism.

The Copycat system (Hofstadter and Mitchell 1994) has a somewhat similar feedback mechanism as our STM. Its global 'temperature' and the 'unhappiness' of objects serve as measures controlling the random choices that facilitate unexpected behavior. The main differences are that it works on textual objects instead of continuous signals, and its architecture is based on a crowd of heterogeneous codelets instead of neural networks. The latter feature makes it more reminiscent to Brooks' robots.

Relaxation of hard constraints, e.g. leaving the physical space and thinking in another context by analogy or metaphor, would call for higher level conceptual models than neural networks, and is out of the scope of this paper. The same applies to reflective thinking. Our poor system itself does not recognize creativity, though it may be possible to detect it from the abrupt changes happening in the STM values during a creative leap.

Had the system a measure of cumulative effort used before a successful trial, or about the time spent without a goal at all, it could model the emotional frustration and boredom that are supposed to control creative behavior on a higher level. In previous work (Takala 2005) these were used to control the recruitment of alternative methods to solve given problems. Combining the mechanisms with the present work may result in interesting behaviors.

Our general approach follows much that suggested in robotics (Brooks 1991, Brooks and Stein 1994). Although the current implementation is based on a single neural network, and a multilevel hierarchical organization of several SOMs may be possible, a more heterogeneous architecture may also be due.

Conclusion

This work emphasizes the contextual nature of creativity, culminating to expectations and their role as soft constraints that must be violated in order to find novel and surprising solutions to problems. Concentrating on the preconceptual level of cognition, it contributes to an area rarely touched in previous works.

A computational model is presented that implements a primitive form of creativity, which may serve as a basis for further development. Autonomous formation of conceptual spaces is demonstrated with the self-organizing memory, and a learning mechanism proposed that simulates the temporary preferences typical in idea fixations. Though our example case is about kinematics, the model is domain independent and may be applied in many different areas.

The creativity model proposed in this paper is based on various ideas that are not novel as such but presented in multiple previous works. The main contribution appears to be the implementation where a self-organizing neural network is combined with control mechanisms usually applied on the symbolic level. Our system is not using predefined heuristics or encoded algorithms but applies generic learning principles to form (pre)concepts, on which the feedback mechanism operates.

A theoretical conclusion is that creativity cannot happen just anywhere, but requires certain conditions: In order to be surprising, the situation should involve expectations, or temporary preferences, that are violated in a creative act. If the system acts in a continuous parametric domain, such as movement, the setting (or its representation) should be non-monotonic, such that the system may get stuck in a local optimum. Yet another condition, though mostly overlooked in the present work, is motivation. If the problems to be solved are given from outside, the system acts in a slave mode, whereas a truly creative mind would be curious and willing to set problems, not just to solve them.

An immediate future work is to study the proposed mechanism in more complicated cases, such as a real robot, taking into account physical continuity of movement and not only static positions. Another extension is to facilitate explorative creativity by letting the robot move randomly around and learn continuously. Long term goals include developing the proposed approach towards higher-level cognition and conceptual thinking, including analogical reasoning and emotional self-control.

Acknowledgements

Thanks to the anonymous reviewers for their comments on the manuscript and pointers to additional sources. This work has been made possible in part by the sabbatical system of Aalto University.

References

Boden, M. (1990). *The Creative Mind: Myths and Mechanisms*, Weidenfeld & Nicholson (2nd ed. Routledge, 2004).

Brooks, R. (1991). Intelligence without representation. *Artificial intelligence, 47*(1), 139-159.

Brooks, R., & Stein, L. (1994). Building brains for bodies. *Autonomous Robots, 1*(1), 7-25.

Chella, A., Gaglio, S., Oliveri, G., Augello, A., & Pilato, G. (2014). Creativity in Conceptual Spaces. In *5th Intl. Conf. on Computational Creativity*, 306-314.

Csikszentmihalyi, M. (1997). *Creativity: flow and the psychology of discovery and invention*. Harper Perennial.

Duka, A. V. (2014). Neural network based inverse kinematics solution for trajectory tracking of a robotic arm. *Procedia Technology*, *12*, 20-27.

Egner, T., Monti, J. & Summerfield, C. (2010). Expectation and Surprise Determine Neural Population Responses in the Ventral Visual Stream. *The Journal of Neuroscience, 30(49)*, 16601–16608.

Gabora, L. (2010). Revenge of the 'neurds': Characterizing creative thought in terms of the structure and dynamics of human memory. *Creativity Research Journal, 22*(1), 1-13.

Gärdenfors, P. (2004). *Conceptual spaces: The geometry of thought*. MIT press.

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78.

Grace, K., & Maher, M. L. (2014). What to expect when you're expecting: the role of unexpectedness in computationally evaluating creativity. In *4th Intl. Conf. on Computational Creativity*, 120-128.

Hofstadter, D., & Mitchell, M. (1994). The copycat project: A model of mental fluidity and analogy-making. *Advances in connectionist and neural computation theory, 2*(31-112), 29-30.

Hristovski, R., Davids, K., Araujo, D., & Passos, P. (2011). Constraints-induced emergence of functional novelty in complex neurobiological systems: a basis for creativity in sport. *Nonlinear Dynamics, Psychology and Life Sciences*, 15(2), 175-206.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design studies*, *12*(1), 3-11.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological bulletin*, *70*(6p1), 387-403.

Kohonen, T. (2001). *Self-organizing maps* (Vol. 30). Springer Science & Business Media.

Kohonen, T., Kaski, S., Lagus, K., Salojarvi, J., Honkela, J., Paatero, V., & Saarela, A. (2000). Self organization of a massive document collection. *Neural Networks, IEEE Transactions on*, *11*(3), 574-585.

Lorini, E., & Falcone, R. (2005). Modeling expectations in cognitive agents. In *AAAI 2005 Fall Symposium: From Reactive to Anticipatory Cognitive Embodied Systems*. 114-121.

Maher, M. L., Brady, K., & Fisher, D. (2013). Computational models of surprise in evaluating creative design. In *4th Intl. Conf. on Computational Creativity*, 147-151.

Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creativity Research Journal*, *15*(2-3), 107-120.

Newell, A., Shaw, C., & Simon, H. A. (1959). *The processes of creative thinking*. Santa Monica, CA: Rand Corporation.

Pylyshyn, Z. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition* 80(1), 127-158.

Rosch, E. (1973). Natural categories. *Cognitive psychology*, 4(3), 328-350.

Rosch, E., Thompson, E., & Varela, F. (1992). *The embodied mind: Cognitive science and human experience*. MIT press.

Stewart, J. R., Gapenne, O., & Di Paolo, E. A. (Eds.). (2010). *Enaction: Toward a new paradigm for cognitive science*. MIT Press.

Takala, T. (2005). Creativity as disruptive adaptation – a computational case study. In: Gero and Maher (eds.), *Computational and Cognitive Models of Creative Design VI*, Univ. of Sydney, 95-110.

Wiley, D. J., & Hahn, J. K. (1997). Interpolation synthesis of articulated figure motion. *Computer Graphics and Applications, IEEE*, *17*(6), 39-45.