Procedia Computer Science
Volume 88, 2016, Pages 21–26
7th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2016

Pattern Turnover within Synaptically Perturbed Neural Systems

Stephen L. Thaler
Imagination Engines, Inc., St. Charles, Missouri, U.S.A.

Abstract
A critical level of synaptic perturbation within a trained, artificial neural system induces the nucleation of novel activation patterns, many of which could qualify as viable ideas or action plans. In building massively parallel connectionist architectures requiring myriad, coupled neural modules driven to ideate in this manner, the need has arisen to shift the attention of computational critics to only those portions of the neural "real estate" generating sufficiently novel activation patterns. The search for a suitable affordance to guide such attention has revealed that the rhythm of pattern generation by synaptically perturbed neural nets is a quantitative indicator of the novelty of their conceptual output, that cadence in turn characterized by a frequency and a corresponding temporal clustering that is discernible through fractal dimension. Anticipating that synaptic fluctuations are tantamount in effect to volume neurotransmitter release within cortex, a novel theory of both cognition and consciousness arises that is reliant upon the rate of transitions within cortical activation topologies.

Keywords: artificial consciousness, artificial neural systems, computational creativity, neural activation topologies, pattern turnover, volume neurotransmitter release

1 Introduction

A patented neural architecture called a "Creativity Machine" (Thaler, 1996a, 1997a, 1997b, 2013, 2014) utilizes an artificial neural system (ANS) whose inputs are pinned within some environmental context as its connection weights stochastically fluctuate to produce a succession of output patterns (Thaler, 1995). A computational critic, typically taking the form of another neural system, associates each of these emerging patterns with an anticipated pattern-based consequence whose scalar distance, δ, from some goal pattern is used to modulate the amplitude of the aforesaid synaptic fluctuations. Allowed to run autonomously, such a neural system may eventually arrive at a solution pattern, at which time the δ metric approaches zero and the synaptic perturbations proportional to it subside (Thaler, 1998).

After numerous engineering successes in applying this neural paradigm in the area of autonomous discovery and invention (Thaler, 1997b, 2013), research has focused on the statistical mechanics of such

Selection and peer-review under responsibility of the Scientific Programme Committee of BICA 2016

doi:10.1016/j.procs.2016.07.401

systems, namely the relationship between synaptic fluctuation level and the novelty of the candidate patterns emerging from any connectively perturbed neural network, referred to herein as an "imagitron." The mean level of ongoing synaptic fluctuations, $\langle \delta w \rangle$, among the net's N_s connection weights governs its pattern evolution. At any moment, we may represent this perturbation level by

$$\langle \delta w \rangle = \frac{\sum_{i=1}^{N_s} (w_i^{'} - w_i)}{N_s}, \qquad (1)$$

with the term $w_i' - w_i$ expressing the instantaneous difference between the perturbed and unperturbed values of the i^{th} connection weight. Should all such terms be of the same sign (i.e., exclusively excitatory or inhibitory perturbations), the net's pattern turnover as a function of mean synaptic perturbation, $\langle \delta w \rangle$, takes the form shown above in Figure 1, with the dotted curve reflecting overall pattern turnover (i.e., memories and confabulations combined) from the net. Memory turnover, shown via the solid curve, typically peaks at a critical level, $\langle \delta w \rangle_c$, that separates regimes of intact memory output (U) and mild confabulation generation (V). Higher levels of fluctuation in synaptic integration in the W regime typically result in purely confabulatory output often amounting to nonsense.

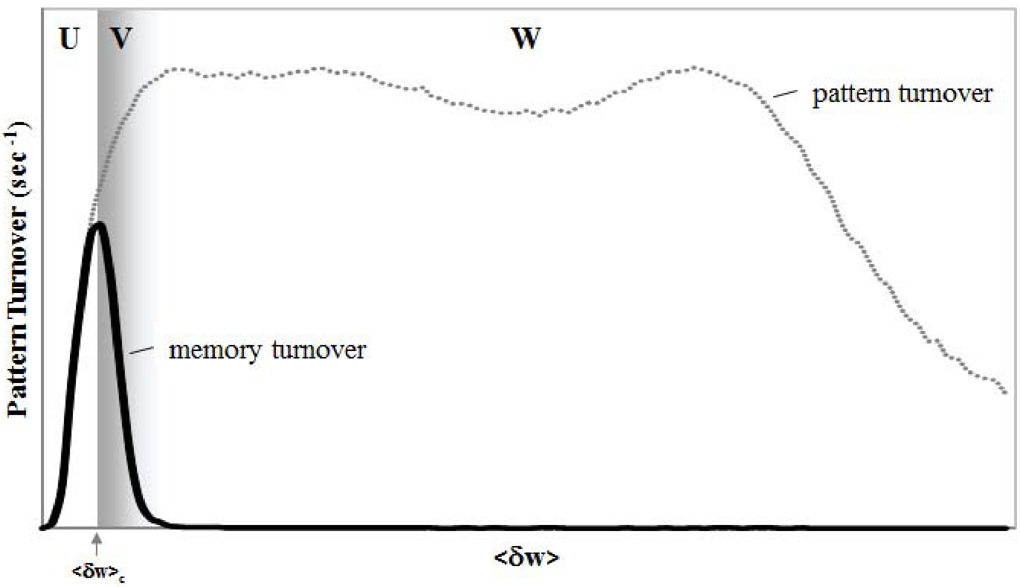

Figure 1. Pattern turnover of a synaptically perturbed ANS with pinned inputs. At the critical threshold, $\langle \delta w \rangle_c$, novel patterns begin to dominate system output.

Important to note is that the rhythm of pattern generation varies during the progression from low to high fluctuation levels in synaptic perturbation. Within the U regime, pattern turnover, consisting primarily of memories, increases in frequency from zero at $\langle \delta w \rangle = 0$, becoming less sporadic as perturbation level increases through the U regime. At $\langle \delta w \rangle_c$, pattern turnover, still consisting largely of memories, becomes rapid and linear. Within the V regime, where confabulatory patterns begin to dominate network output, the rhythm of memory generation decreases and becomes more tentative. Thereafter, in the W regime, overall pattern turnover generally plateaus and then declines, becoming more sporadic with increasing synaptic fluctuation, until finally approaching zero at sufficient perturbation level when saturating synaptic integration pins all neurons at a fixed activation state.

From the standpoint of implementing computational creativity within artificial neural systems, operation at or slightly above the steady-state perturbation level, $<\delta w>_c$, provides novel activation patterns that only slightly depart from memories stored within the neural system's attractor landscape and as such constitute plausible notions that largely obey the learned constraints of the absorbed conceptual space. Accordingly, Creativity Machines (CM) operate at or near this synaptic fluctuation level either by design or through their own self-governance via feedback mechanisms between critics and imagitrons. Complex CMs incorporating multiple imagitrons manifest the same behavior as depicted in Figure 1, as these neural modules interconnect via fluctuating synaptic connections, such disturbances driving the formation of semi-plausible combinations of memories and confabulations into compound concepts (Thaler, 1996b). However, as these systems become more extensive, perceptrons or other critic algorithms must first locate such novel ideational juxtapositions prior to evaluating them, choosing then to either log the resulting discoveries, implement them as strategies, and/or reinforce them as memories.

Motivated by the need to locate such notions as they are forming within very large-scale neural systems, attention has centered on determining the optimal affordances to use in locating potential ideas as they activate. The experiment described below is both a summary of and an expansion upon prior work along these lines (Thaler, 2014), its conclusions forming the basis of a novel approach to both computational creativity and machine consciousness.

2 Assessing Novelty from Neurodynamics

The goal of this research was to determine the novelty of patterns emerging from a synaptically perturbed artificial neural net (ANN) based solely upon peripheral clues such as the time evolution of neural activation patterns. Finding such observables would fulfill the engineering objective of detecting idea formation among vast ensembles of chaotic neural nets.

The parameters monitored in the course of these experiments included (1) the number of perturbations, N_p, randomly hopping between the ANN's connection weights, (2) the fixed or average magnitude of such synaptic perturbations, σ, (3) the total number of neurons in the network, N_n, and (4) the number of cycles, N, of random placements of such perturbations required to produce N_0 output pattern transitions. Accordingly, the quantity $N_p\sigma/N_n$ represented the ensemble-averaged, perturbation-induced bias in synaptic integration per neuron, that integration executing a random walk, in steps proportional to integer multiples of σ at each individual processing unit to drive the net's pattern turnover.

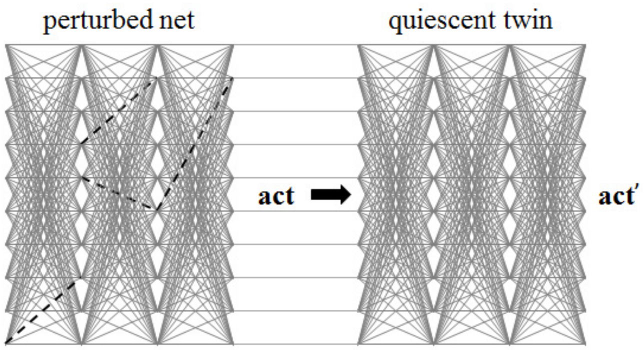

Figure 2. Pattern turnover of synaptically perturbed ANN filtered by its quiescent twin. Dashed connections represent transient synaptic perturbations.

In each of these computational experiments, the perturbed net took the form of a pre-trained, auto-associative multilayer perceptron. The unperturbed twin of this chaotic net served to filter the stream of perturbation-driven output patterns for their novelty (Figure 2) based upon the root-mean-square error, Δ^*_M**act**, between any of the perturbed net's output patterns, **act**, and the quiescent net's resulting output pattern, **act′**. This reconstruction metric measured how closely any given noise-driven output pattern, **act**, came to any of either net's stored memories. Allowed to pass only those perturbation-driven patterns falling within some narrow range of reconstruction errors**,** the latter net could isolate pattern streams having constant novelty. Serving as a measure of rhythmic linearity, fractal dimension, D_0, was analytically determined via Mandelbrot measures (Peitgen and Saupe, 1988) for each of these isolated "novelty bands."

Monitoring the perturbed net's stream of output patterns using its quiescent clone, it was empirically established that for a given novelty band, $\Delta^*_M\mathbf{act} \pm \delta(\Delta^*_M\mathbf{act})$, the relationship between fractal dimension D_0 of the output stream (i.e., linearity), the required number of perturbation cycles, N to produce N_0 successively distinct patterns (i.e., frequency), and non-vanishing novelty, $\Delta^*_M\mathbf{act}$, was given by

$$D_0 \approx \frac{\ln\left(\kappa N / \Delta^*_M\mathbf{act}\right)}{\ln N}, \tag{2}$$

with κ constant (Thaler, 2014) for a given choice of N_p, σ, and N_n. Confirmation of this equation came from the theory of fractional Brownian motion, fBm, (Mandelbrot and Van Ness, 1968). In addition to substantiating Equation 2, this random walk model served to identify the numerator on its right-hand side as the logarithm of the number of excursions, N_0, of the perturbation-driven output pattern to the vicinity of what is at any moment its closest stored memory within the net.

Noting that fractal dimension, D_0, of a temporal stream may vary from 0, indicating an isolated event, to 1, representative of a linear sequence, the interpretation of Equation 2 becomes intuitively appealing. In essence, this relationship shows that there is an intrinsic tradeoff between the hesitancy of pattern delivery, reflected in D_0, and the frequency inherent within the number of perturbation cycles, N, required to generate $N_0 = \kappa N / \Delta^*_M\mathbf{act}$ output patterns. Further analysis showed that κ was the average step size in the random walk of synaptic integration at each neuron, approximately $N_p\sigma/N_n$. From this perspective it was then determined that near the critical point $\langle\delta w\rangle_c$,

$$D_0 \approx \frac{1}{2} + \frac{\frac{1}{2}\ln\left(\frac{N_p}{N_n - N_p}\right)}{\ln(N)}. \tag{3}$$

According to this relationship, fractal dimension of the pattern stream approaches a value of ½ when half the net's neurons are populated on the average by a single synaptic perturbation (i.e., $N_p = N_n/2$).

The results of separate cognitive experiments obeyed the same relationship expressed by Equation 2 for articulated stream of consciousness within a population of human volunteers engaged in divergent creative exercises requiring various degrees of originality (Thaler, 2014). Qualitatively, as the level of novelty required by the task increased, the lower were both the measured frequency (equivalent to 1/N) and fractal dimension (D_0) of the ideational stream. More importantly, the value D_0 = ½ appeared to be the boundary between rote memory recall and original thought, suggesting through Equation 3 that within the brain, it is possible that perturbative pinning of neurons begins as the number of synaptic perturbations, N_p, begins to approach the number of neurons, N_n. Such synaptic saturation thusly impedes the internal pattern completion process that generates intact memories, leading instead to the nucleation of mild and potentially useful confabulations upon synaptic noise.

3 Discussion

Filtering the pattern turnover of a synaptically perturbed neural net on the basis of novelty reveals two readily observable metrics that are useful in determining how far output patterns are wandering from the neural system's direct experience, the first being the frequency of pattern output, and the second, the non-linearity of such pattern turnover. These affordances relate respectively to N, the number of perturbation cycles required to generate a targeted number of output pattern transitions, and the fractal dimension, D_0, essentially a measure of the temporal clustering of these transitions.

Therefore, in compound CMs consisting of multiple neural modules globally suffused with synaptic perturbations, an efficient way of detecting novel pattern formation is through a continuous anomaly scan of the neural environment in search of the more tentative (i.e., $D_0 \leq ½$) and lower frequency activation streams. With their attention thusly drawn to these novelties, critics may now assess their utility or value, choosing then to modulate synaptic perturbation level into either the U or W regimes where Hebbian learning within a quiescent environment may reinforce these notions into memories (Thaler, 2013, 2014).

4 Connecting with Neurobiology

The model yielding Equations 2 and 3 assumes only a system of switching elements, each functioning via weighted integration of input signals and threshold firing behavior, the entire system flooded by weight disturbances that drive an ongoing stream of activation patterns. Realizing that neurobiology fulfills these same minimal requirements, similar pattern delivery rhythms should be occurring within the brain, where the equivalent of random synaptic perturbation would be volume-released neurotransmitters providing the necessary fluctuations to normal synaptic transmission, as well as fine control over whether cortex is in a rote or creative mode. Accordingly, the pace of human ideation has shown the same characteristic behaviors. It may be fast and linear during rote memory recall, or slow and tentative when creatively engaged, strongly suggesting that some form of novelty filter, similar to that shown in Figure 2, could be at work in the brain, constantly habituating to cortical state to reveal new or anomalous activation patterns forming therein. Temporal distribution algorithms could also be incorporated within the brain's cognitive algorithms, being either explicit, estimating temporal distribution (e.g., fractal dimension) of pattern streams, or implicit, relying upon adaptive novelty detectors that would naturally react to the rarer and more sporadic neural activations related to idea formation.

Whereas the theory of pattern turnover proposed herein serves as a first order approximation to human cognition, certain engineering challenges arise when attempting to extend the model to the hundreds of billions of neurons characteristic of the human cortex. There the novelty detection system would require a prohibitively large number of input nodes and an even larger number of connection weights to absorb the entire cortical status quo. In overcoming such scaling issues, a patent-pending methodology, based upon the rhythm of pattern delivery, has been developed to monitor the myriad pattern streams simultaneously emerging from very large cortical simulations. This approach has in turn inspired a cognitive model (Figure 3) based upon the noise-driven formation of memory chains activating through momentary unions of multiple neural modules to create compound ideas that in turn recruit associative chains likewise composed of linked neural modules that collectively encode emotional response to freshly forming notions. These ideational chains (e.g. ABC), combined with affective responses (e.g., DEFG), would be observed via descending connections from cortex to thalamus within which a novelty-filtered and condensed version of the cortical chain A-G forms (e.g., abcdefg). Such a topologically faithful thalamic digest would synapse with neural modules such as H, specialized at flooding cortex with neurotransmitters that either excite or quench activation turnover.

Crucial to both the proposed synthetic, cognitive architecture and biological consciousness would be the detection of idea formation among myriad neural modules transiently interconnecting into notions and the accompanying subjective responses to them. Fortunately, the same mathematical relationships governing pattern turnover hold in the topologically based system alluded to in Figure 3, since the affective responses are driven at the same characteristic rhythms as the ideations to which they are transiently binding. Thus, such conceptual chains, now synchronized with their emotional responses, are detectable through their activation frequency or fractal dimension. Of course, such a filtering process would need to be sensitive to a specified frequency band in the V regime, so as not to capture memories or gibberish, but the mild confabulations we call ideas.

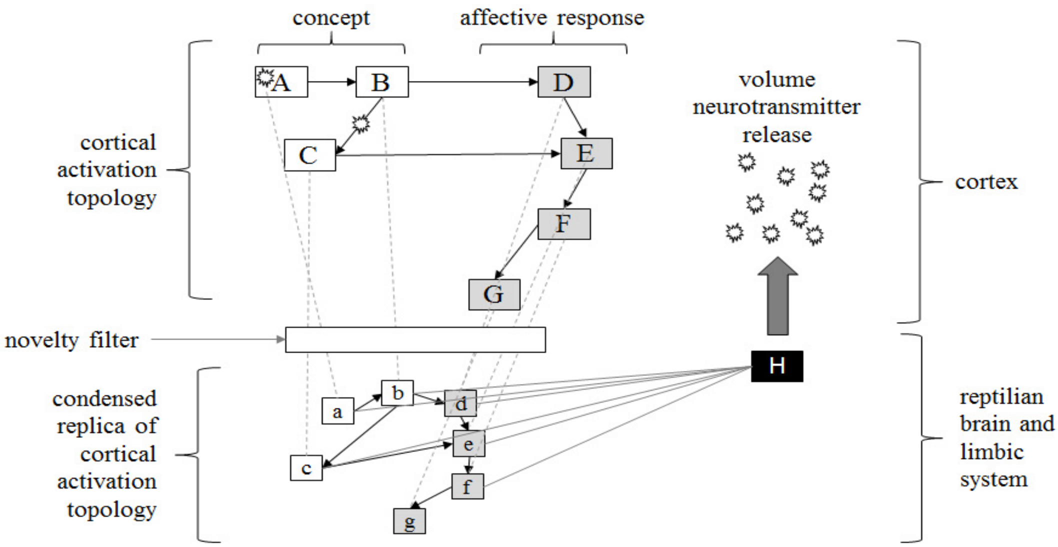

Figure 3. Brain viewed as compound Creativity Machine. (Machine implementation patent pending.)

References

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motion, fractal noises and applications, *SIAM Review* **10**(4):422-437.

Peitgen, H. and Saupe, D. (Eds.) (1988). *The Science of Fractal Images*, New York, NY: Springer-Verlag.

Thaler, S.L. (1995). "Virtual input phenomena" within the death of a simple pattern associator. *Neural Networks*, **8**(1):55-65.

Thaler, S.L. (1996a). Network "cavitation" in the modeling of consciousness, presented at *Toward a Science of Consciousness*, Tucson, AZ, March, 1996.

Thaler, S.L. (1996b). A proposed symbolism for network-implemented discovery processes, presented at *World Congress on Neural Networks*, (WCNN'96), San Diego, CA, 1265-1268, Mahwah, NJ: L. Erlbaum.

Thaler, S.L. (1997a). U.S. Patent No. 5,659,666, Washington, DC: U.S. Patent and Trademark Office.

Thaler, S.L. (1997b). A quantitative model of seminal cognition: the creativity machine paradigm, presented at *Mind II Conference*, Dublin, Ireland, 1997.

Thaler, S.L. (1998). Predicting ultra-hard binary compounds via cascaded auto- and hetero-associative neural networks. *Journal of Alloys and Compounds*, **279**(1998): 47-59.

Thaler, S.L. (2013). The creativity machine paradigm, In E.G. Carayannis (Ed.), *Encyclopedia of Creativity, Invention, Innovation, and Entrepreneurship*. (pp. 447-456). New York, NY: Springer Science+Business Media, LLC.

Thaler, S.L. (2014). Synaptic perturbation and consciousness. *International Journal of Machine Consciousness*, **6**(2):75-107.