Thinking in Prose and Poetry: A Semantic Neural Model

Sarjoun Doumit, Nagendra Marupaka and Ali A. Minai, *Senior Member, IEEE*

***Abstract*— The neural basis of creative thinking – indeed of all thinking – remains mysterious. One influential theory by Mednick holds that creative thinking reflects a difference in the associational structure of conceptual representations in the mind. We have previously proposed a neural network model based on itinerant dynamics to model thinking, and used it to show that a small-world, scale-free associational structure – similar to that found empirically in linguistic data – is especially efficient for exploring conceptual space and generating conceptual combinations. In this paper, we apply this model to associative networks obtained from the poetry of Dylan Thomas and John Gay, and the prose of F. Scott Fitzgerald and George Orwell. Network analysis shows that poetic texts indeed incorporate a wider distribution of associations than prose. However, neural simulations using semantic networks from the four sources present a more complex picture. We also consider the case where a poet's associative network is transformed to that of a prose-writer to test the impact of this manipulation.**

I. Introduction

Most models of creativity postulate that new ideas emerge through the combination of concepts found in existing ideas, and that the creative process is primarily one of *conceptual combinaton* with unusual combinations being potentially creative [1], [2], [3], [4], [5]. This view is also supported by the experience of creative thinkers such as Einstein and Poincaré [6], [3].

The work in this paper is motivated by an influential proposal by Mednick [6] that the difference between creative and non-creative thinkers arises from the distribution of conceptual associations in their minds. The suggestion is that non-creative individuals usually have only a few conventional associations with every concept while creative individuals have a broader set of associations, including unusual – or "remote" – ones. Mednick described the difference as that between a steep or flat associative hierarchy. The principle underlying this view is that thought is fundamentally an associative phenomenon, where new ideas emerge because they are triggered associatively by current ideas or external primes [7], [2].

From a networks perspective, the implication of Mednick's postulate is that the associative semantic networks of creative and non-creative individuals have different connectivity structure, the latter showing a "fatter" tail than the former. Recently, we proposed a simple neurodynamical model to simulate associative conceptual combinations, and used it to look at the features of the emergent dynamics in several canonical semantic network models [8], [9]. We found that, under the model, semantic networks with a power-law small-world connectivity were the most productive and efficient at producing coherent conceptual combinations. In fact, this is precisely the type of connectivity found emprically in many studies of associations in texts and experimental data [10], [11], [12], [13]. However, the networks we simulated were abstract, and even the data used in empirical studies is usually from structured sources such as dictionaries, thesaurii and association tables rather than fluent thought. As a way to address this issue in a preliminary way, we have extracted associative networks from four textual sources with varying conventionally accepted levels of creativity, and simulated our neural model based on these with the goal of discerning signature phenomena. Preliminary results from this study are presented here.

Sarjoun Doumit (Email: doumitss@mail.uc.edu), Nagendra Marupaka (Email: marupana@mail.uc.edu), and Ali Minai (Email: Ali.Minai@uc.edu) are all with the School of Electronic and Computer Systems, University of Cincinnati, Cincinnati OH 45221

Acknowledgement: This work was supported in part by a National Science Foundation CreativeIT grant to Ali Minai (IIS-0855714) and a National Science Foundation INSPIRE grant to Ali Minai (BCS-1247971).

II. Background and Motivation

Since the pioneering work of Guilford [14], many theoretical [1], [6], [2], [3], [4], [15], [16], [17], [18], [5] and computational [7], [2], [15], [19], [17], [18], [20], [21] approaches to creativity have been proposed. Creativity has also been studied experimentally in the context of brainstorming [22], [23], [24], [25], [26], [27], [28], [16] and through brain imaging [29], [30]. However, the neural basis of creativity - and thought in general - remains mysterious. An important observation in this regard came from Hebb [31], who postulated that novel combinations of ideas might result from the exploratory combinatorial firing of neural assemblies in the brain. This broad idea was developed by Thagard and Stewart [21] into an explicitly neural model based on the theory of neural binding through convolution. We have recently proposed a broadly similar but more comprehensive model based on modulated itinerant dynamics in recurrent neural networks [32], [33], [34], [20], [35]. The system used in the current paper is a component of this larger model [8], [9], and is used here with real-world associative semantic networks to explore issues of semantic connectivity and creativity.

The overall approach we take is based on two postulates:

1. New ideas are generated by the selective recombination of existing ideas in the minds of individuals [1], [6], [36], [37], [23], [38], [26], [3], [5], which can be seen as dynamic cognitive-epistemic networks – henceforth termed *epistemic networks* for brevity.
2. The creativity of individuals depends on the organization of knowledge in their epistemic networks [6], [39], [8], [9].

The method we use is: a) Extracting epistemic networks from artifacts of thought, such as writings by specific authors, and comparing them to determine any systematic characteristics – especially with regard to creativity, and b) Using neurodynamical simulations based on the extracted networks to explore the *generative potential* of the networks with their specific connectivities – in essence, to ask the question: What types of trains of thought do different types of minds generate?

III. Modeling Epistemic Networks

Associative memory has been studied experimentally mainly using cued recall and word associations [40], [41]. There are also several empirical studies of associative semantic networks derived from the association norms as well as thesausrii, dictionaries, databases and ontologies [10], [11], [12], [13], [39]. Recently, this approach has been applied rigorously to associations collected through a sequential process that is somewhat closer to a "train of thought" [42]. While extremely useful, none of these necessarily captures the actual pattern of associations that might come into play when people think or write *fluently* in natural settings. Indeed, obtaining reliable experimental data on such associations is likely to be quite difficult, but is critical if thought is to be understood properly in its "natural habitat".

One possible way to get some limited access to a mind's associative structure – in a very approximate sense – is by looking at the concrete artifacts produced by fluent thought, such as writings or speeches. While the best way to do this remains uncertain, we make a preliminary attempt to explore this approach by extracting associations based on three intuitive assumptions:

1. Fluent thinking during authorial composition is organized into sequential *semantic units* (SUs), each of which can be regarded as a coherent semantic entity.
2. The joint appearance of two concepts in the same semantic unit indicates an association between them in the author's mind.
3. The strength of association between two concepts in the author's mind can be estimated by their probability of co-occurrence in a semantic unit over a significant corpus of the author's work.

Of course, these postulates are extremely simple – even simplistic – but they provide a way to begin looking at texts as a direct window into minds without obscuring it with artificial laboratory protocols or the machinery of natural language processing and ontology construction. In the same spirit, we make the practical choice of defining semantic units as sentences or stanzas and concepts as words. We ignore the order of occurrence for words within an SU. All these choices can be revised in later work.

Using the approach described above, we extracted epistemic networks from four sources:

1) Ninety-nine poems by Dylan Thomas (20th century surrealist English poetry).
2) Thirty-six poems by John Gay (18th century Augustan English poetry).
3) *The Great Gatsby* by F. Scott Fitzgerald (20th century American fiction book)
4) *Politics and the English Language* by George Orwell (20th century English non-fiction essay).

The connectivity of the epistemic network for author q is represented by the $n_q \times n_q$ *association matrix*, A^q , where n_q is the number of distinct words in the corpus for q, and element a^q_{ij} of A^q represents the fraction of SUs in which both words b^q_i and b^q_j occur. We call a^q_{ij} the *associative weight* between i and j. Defining the semantic unit is a complex issue for poets – especially with those like Dylan Thomas who use punctuation irregularly. We made the simple decision to segment SUs by periods, exclamation points, question marks and blank lines (indicating end of stanza). This means that SUs for poetic corpora are highly variable – ranging from short sentences through multi-line stanzas in a few cases. For example, the following is one SU from Gay:

But I, who ne'er was bless'd by fortune's hand,
Nor brighten'd plough shares in paternal land,
Long in the noisy town have been immur'd,
Respir'd its smoke, and all its cares endur'd,
Where news and politics divide mankind,
And schemes of state involve the uneasy mind:
Faction embroils the world; and every tongue
Is mov'd by flattery, or with scandal hung:
Friendship, for sylvan shades, the palace flies,
Where all must yield to interest's dearer ties,
Each rival Machiavel with envy burns,
And honesty forsakes them all by turns;
While calumny upon each party's thrown,
Which both promote, and both alike disown.
(Rural Sports: A Georgic - Canto I)

However, in the absence of a clearly better alternative, we follow the convention described above. We base this choice on the argument that the variability in semantic unit size is an essential aspect of the creativity involved in poetry – that the poet thinks in more complex and connected "web-like" units that turn into complex sentences and stanzas, while the prose writer – even in fiction – is constrained by the linear structure of narrative or description.

Stop words such as '"the", "I", "your", "of", etc., are removed from the text and not included in the co-occurrence matrix. Given the limited size of each corpus, many words only occurr a few times in the whole corpus, and all analysis in simulations in this paper use just the 500 most frequent words for each author. Thus, $n_q = 500 \equiv N$ in all cases. The words (and the A matrix rows and columns) are sorted in order of decreasing frequency for all cases. Figures 1 and 2 show the sub-networks obtained for Thomas and Fitzgerald for the top 100 words in each (showing all 500 words makes

the networks hard to see).

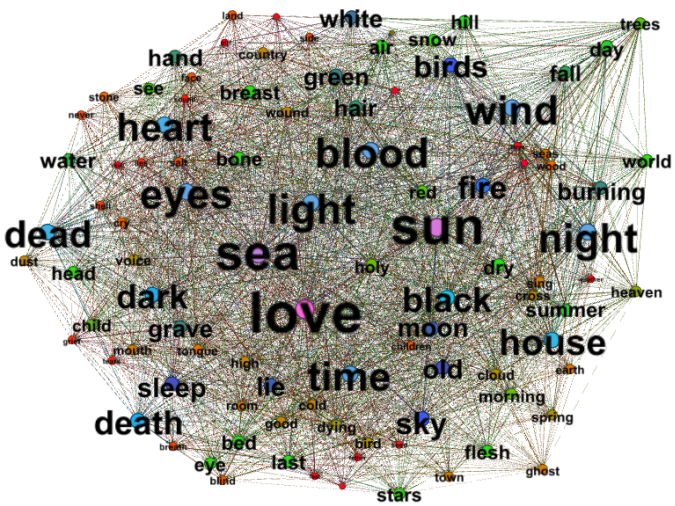

Fig. 1. The network of the 100 most frequent words for Thomas. Word size reflects frequency of use and node colors identify communities obtained through the algorithm in [43]

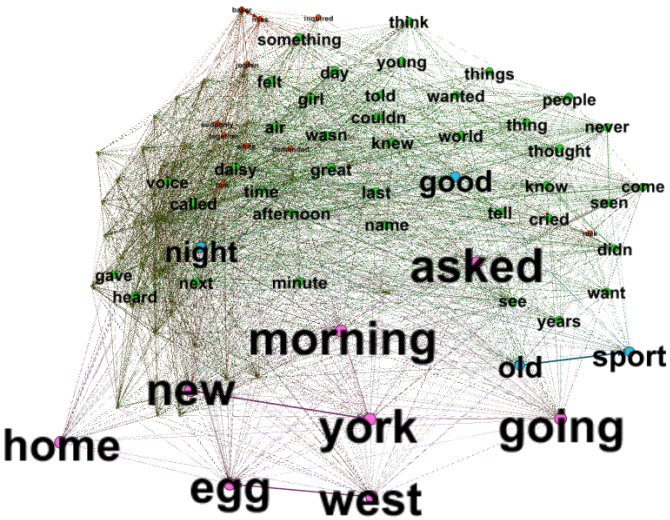

Fig. 2. The network of the 100 most frequent words for Fitzgerald. Word size reflects frequency of use and node colors identify communities obtained through the algorithm in [43]

The basic statistics of the four corpora are shown in Table I. While they differ in several ways, the first three (Thomas, Gay and Fitzgerald) have more than $12,000$ words and vocabularies of roughly comparable sizes. The Orwell corpus is much smaller, and therefore atypical. Nevertheless, it is included as a representative of non-fiction writing.

IV. Network Analysis

Each of the networks was analyzed for various network attributes, which are summarized in Table II. All networks are connected but sparse, with the Orwell network sparsest of all (mean degree 17.85). The *mean shortest path length* (MSPL) for the Thomas, Gay and Fitzgerald networks is very

Author	Number of Items	Number of Semantic Units	Total Words	Vocabulary Size
Thomas	99	846	12,765	4,790
Gay	36	883	15,901	5,577
Fitzgerald	1	3,253	18,233	5,417
Orwell	1	207	2,019	1,277

TABLE I
NETWORK STATISTICS

small, whereas that of the Orwell network is significantly larger. The *clustering coefficient* for a node is defined as the fraction of its neighbors that are also directly connected [44], and the clustering coefficient for the network is taken to be the mean over all nodes. All networks show significant clustering, but the Orwell network is much more clustered than the other three. *Modularity* represents the degree to which a network can be seen as a set of modules with strong internal connectivity and weak external connectivity [43]. The Thomas, Gay and Fitzgerald networks have moderate modularity but the Orwell network is highly modular. The *association centrality* of a node is defined as:

$$C_i^q = \sum_{j|a_{ij}^q \neq 0} a_{ij}^q \sum_{k=1}^{N} a_{jk}^q \tag{1}$$

so a node connected with high a-values to nodes with high total a-values is more central. The association centrality of the network is the mean value over all nodes. The networks from the two poets have significantly higher association centrality than those of the non-poets. The centrality of the Fitzgerald network is much lower than the other three, suggesting that most associations are very weak. Finally, the self-similarity metric, $S(A^q)$ is defined as in [45]:

$$s(A) = \sum_{i,j} \kappa_i \kappa_j \tag{2}$$

$$S(A) = \frac{s(A) - s_{min}}{s_{max} - s_{min}} \tag{3}$$

where κ_i is the in-degree of node i, and s_{min} and s_{max} denote the lowest and highest possible values for $s(A)$. The quantity $S(A)$ is a measure of how self-similar the network is at different scales and under different manipulations. It is also a measure of assortativity, i.e., the tendency of nodes with high(low) degrees to connect to other nodes with high(low) degree. All four networks have similar values of S, and are disassortative.

Figure 3 shows the degree distributions for all networks. All authors show an approximately power law plot with an exponential cutoff, though the plot for Orwell is not conclusive. More interestingly, the tails of the distribution for the two poets are similar and fatter than those for the two non-poets. This dichotomy is even clearer in Figure 4, which shows the distribution of mean total associative weights by rank. The associative weights for node j are just the weights a_{ij} sorted in descending order, indicating the

Network	Connection Density	Mean Degree	MSPL	Clustering Coefficient	Modularity	Association Centrality	Self Similarity
Thomas	0.171	85.13	1.83	0.430	0.199	0.0333	0.3345
Gay	0.193	96.18	1.81	0.384	0.183	0.0284	0.2653
Fitzgerald	0.115	57.62	1.90	0.342	0.161	0.0011	0.3021
Orwell	0.036	17.85	2.84	0.692	0.674	0.0157	0.2672

TABLE II
NETWORK ATTRIBUTES

strongest to the weakest associations for the word in node j. These distributions, averaged over all nodes give the *rank-weight distribution* for a network. These distributions provide a partial test of Mednick's hypothesis [6] regarding steep and flat association hierarchies. As can be seen from Figure 4, both the poets – Thomas and Gay – have similar association distributions that are considerably flatter (more fat-tailed) than those for the non-poets. Furthermore, the distribution for the creative prose work (Fitzgerald) lies between those for poetry and non-fiction (Orwell). This partly reflects the fact that poets have semantic units of highly variable size compared to prose authors, but as we argued above, this variability is an essential aspect of poetry. Of course, the small sample of authors used here is insufficient for any conclusive statements, but the results do suggest that creative texts, at least, show flatter association hierarchies than non-creative texts, and poetry more so than prose.

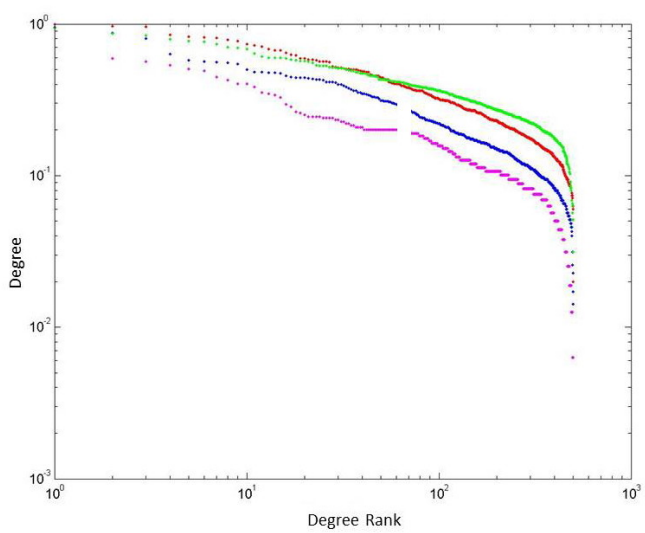

Fig. 3. Rank-degree plots for all four authors: Thomas (red), Gay (green), Fitzgerald (blue), and Orwell (pink). The plot for each author is normalized so that the highest rank node has degree value 1.

The analysis above leads to several salient observations. In particular, it indicates that the poet networks have broader connectivity than the prose networks, making them more "web-like". The Thomas, Gay and Fitzgerald networks all show a strong small-world signature with short MSPL and high clustering [44]. In contrast, the non-fiction network

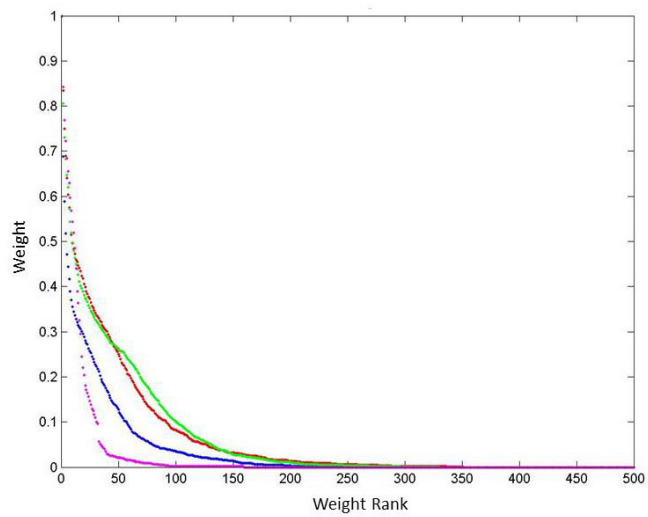

Fig. 4. Normailized rank-weight plot for all four authors

has a chain-like structure with very high local clustering, many communities, and a large MSPL. Though these striking differences for Orwell network may simply reflect the small size of the underlying corpus, it is useful to include it in this study as a concrete example of a network with limited small-world attributes. Another significant observation is that prose networks have far fewer nodes that contribute strongly to the associative structure of the network. Indeed, it was found during analysis (data not shown) that the poet networks had higher associative centrality even for low-frequency nodes, allowing them greater participation in ideas.

V. NEURAL NETWORK MODEL

A. Model Description

Having characterized the extracted associative networks, we instantiated recurrent neural networks using each of them and compared the resulting dynamics of conceptual combination in these networks. The network model used is abstracted from a larger neural model of context-dependent conceptual combination developed recently in our group [34], [20], [35]. We have previously used a version of the model to explore its dynamics for various abstract connectivity patterns [8], [9].

The associative network comprises of N *units*, each representing one word. From a neurobiological viewpoint, each

unit is better seen as an assembly of neurons rather than a single neuron. The connection weights, w_{ij}, between the units are symmetric, and are represented by the *weight matrix*, W obtained from the A matrix by $W = A/max(a_{ij} \in A)$. Thus, the units with the strongest association have a weight of 1, and the rest are scaled proportionately. The net input to a unit i and time t is given by:

$$x_i(t) = \sum_{j=1}^{N} w_{ij}(t)x_j(t) + \gamma_{noise}\xi_i(t) \tag{4}$$

where x_j are outputs from units j, $\xi_i(t)$ is uniform white noise, and γ_{noise}, is a fixed gain. The *state* of unit i at time t is given by:

$$y_j(t) = \alpha y_j(t-1) + (1-\alpha)x_i(t) \tag{5}$$

where α is an inertial parameter typically set to a value just below 1 (we use $\alpha = 0.95$ in all simulations). Thus, the dynamics is a discrete approximation of a continuous one.

Network activity is determined competitively, with the K most highly activated non-refractory units allowed to fire, provided they have $y_i(t) > y_{min}$, where $y_{min} > 0$ is a parameter. The output of unit i is given by:

$$x_i(t) = f(y_i(t)) = \begin{cases} 1, & \text{if } y_i(t) \in \{K \text{ most excited units}\} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

The competitive rule is applied softly so that any unit, j, with $y_j(t)$ within 1% of the nominal cutoff for the K-of-N activity is allowed to fire as well. Thus, even though we use $K = 8$ in all simulations, not all conceptual combinations generated have precisely 8 words.

A key aspect of the model is the use of two types of modulation on the neurons, as described below:

Refractoriness: Each unit i has a *resource*, $r_i(t)$ that depletes when the unit is active and replenishes during inactive periods with the following dynamics:

$$r_i(t) = \begin{cases} (1-\lambda^-)r_i(t-1), & \text{if active} \\ r_i(t-1) + \lambda^+(1 - r_i(t-1)), & \text{if inactive} \end{cases} \tag{7}$$

where λ^- is the *resource depletion rate*, and λ^+ is the *resource recovery rate*. When the resources for i falls below a threshold θ_r^-, the unit becomes refractory and remains so until the resource recovers to a level θ_r^+. The depletion and recovery rates for individual units are not all identical but have a small random variation.

Synaptic Modulation: Experiments have suggested that synapses activated repeatedly over a short period become habituated and temporarily diminish in strength if the activity persists [46]. Once presynaptic activity ceases, they recover gradually to their nominal levels. This is purely a short-term effect independent of any long-term potentiation or depression induced by the activity. In the model, this modulation is represented as follows:

$$w_{ij}(t) = \begin{cases} (1-\psi^-)w_{ij}(t-1), & \text{if active} \\ w_{ij}(t-1) + \psi^+\left[\bar{w}_{ij} - w_{ij}\right] & \text{if inactive} \end{cases} \tag{8}$$

where ψ^- and ψ^+ are, respectively, the synaptic decay and recovery rates, and $\bar{w}_{ij}$ represents the nominal weight of the synapse. Synaptic modulation is an important aspect of short-term memory [47], [48].

B. Model Dynamics

The dynamics of activity in the network is shaped by the competitive process in the context of the underlying connectivity and modulation. If a group of co-active nodes are relatively well-connected with each other, the group can keep winning the competition for activity because of mutual support until forced to shut down by the synaptic or resource modulation – essentially behaving as a *metastable attractor* in the space of concepts (words). In contrast, if the group of co-active nodes are not mutually linked, the group cannot stay co-active and dissipates quickly. Thus, the dynamics of the network is characterized by itinerant transitions between metastable attractors with transient activity in between - similar to the idea of a "mental saccade" [49]. As we have hypothesized before, conceptual combinations corresponding to attractors that persist for a sufficient duration, termed the *awareness threshold*, are perceived consciously as *ideas* [20], [8], [9]. Clearly, the ideas that a network can produce are implicit in the weight matrix of the system, but do not necessarily correspond to ideas that were used explicitly during the encoding/learning process. Many ideas become embedded in the network *implicitly* as a result of the connectivity patterns created by explicit ideas. In one sense, these can be seen as "false memories", but can also be regarded as *emergent ideas* that the mind has become capable of generating as a result of its knowledge. From this perspective, creativity can be regarded as *constructive confabulation* – an observation also made in [19], [50].

Interaction between the two modulation processes plays an important role in shaping the system's dynamics. The synaptic decay rate (0.1 in all simulations) is much faster than the resource decay rate (0.02 in all simulations), so that attractors dissipate on a shorter time scale than the resource depletion in participating neurons. The awareness threshold (set at 20 time-steps in all simulations) is thus determined by the synaptic modulation, and each unit can potentially participate in several ideas (attractors) before becoming refractory. Once depleted, the synaptic strength and resource recover slowly (rate = 0.005), giving other units a chance to participate in ideas. This creates the complex search process for combinations in concept space.

VI. Simulations and Results

Networks of size $N = 500$ were simulated for each of the four authors using identical parameter settings. Simulations were run repeatedly from different initial conditions, each for a duration of $10,000$ time-steps. The competitive threshold was set at $K = 8$, so most generated conceptual

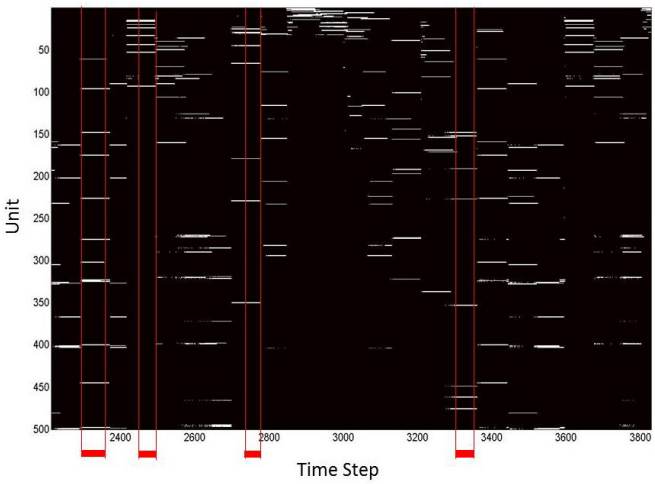

Fig. 5. Network activity over a period when four ideas were generated. Bright lines indicate active units and dark background inactive units. The boundaries of each idea are indicated by vertical lines and marked by the solid rectangles along the x-axis.

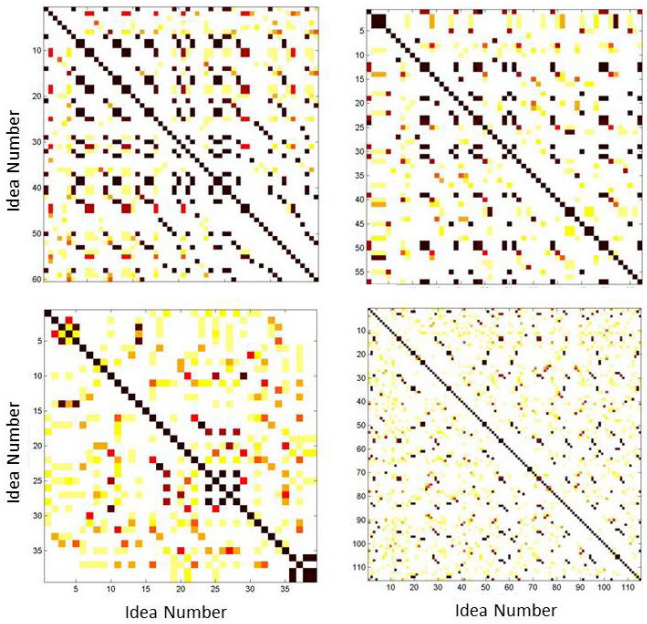

Fig. 6. Idea recurrence plots for one typical run for each of the four authors. Clockwise from top left: Thomas, Gay, Fitzgerald, Orwell.

combinations had 8 concepts – though occasionally as few as 6 and as many as 15 due to soft-thresholding. Figure 5 shows the time-series of activity over a duration when four distinct ideas were generated by the Thomas network. This clearly illustrates the itinerancy between metastable attractors and intervening transient episodes.

Figure 6 shows idea recurrence plots for one typical run from each of the four authors. Each graph shows the pairwise difference (in normalized hamming distance) between all ideas generated sequentially during the run. Lighter colors indicate larger differences. The figure shows that the authors had distinct recurrence patterns. The recurrence pattern for

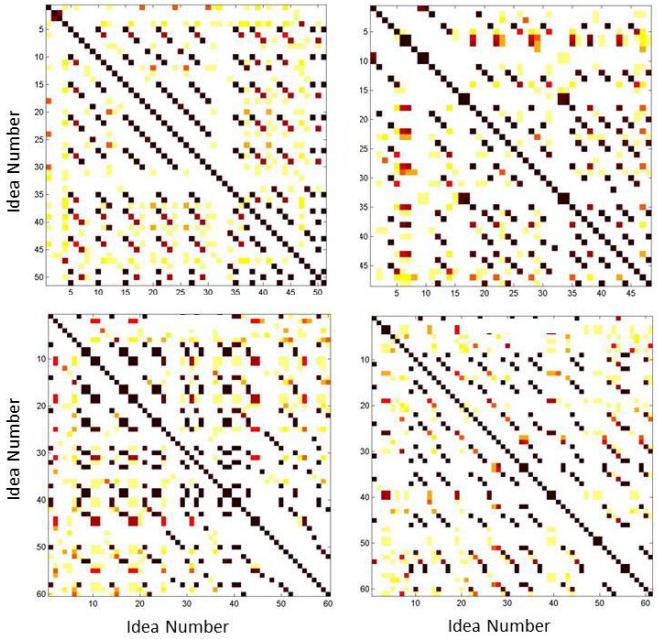

Fig. 7. Four trains of thought for Dylan Thomas

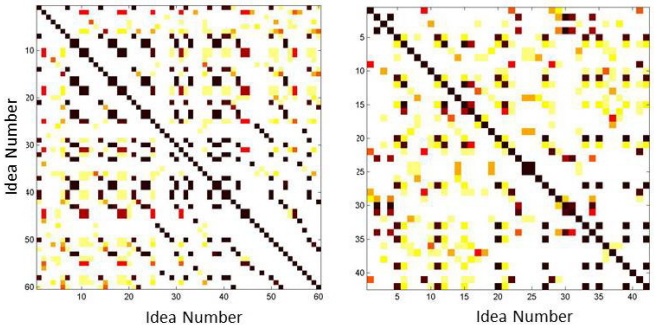

Fig. 8. Recurrence plot for the Thomas network with the weight values of Fitzgerald. Left: Typical Thomas run; Right: Thomas with weights from Fitzgerald.

Thomas indicates exploration in local regions of concept space punctuated by transitions to other regions. The system seems to get into "ruts", representing short sequences of successive patterns that then recur for a period until the dynamics can break free. Overall, the dynamics is characterized by complex recurrences and re-exploration. As a result, a relatively small fraction of ideas in this case (34/60) are unique, but there is a wealth of similar ideas hovering around what may be considered a theme. The plot for Gay shows a similar dynamics, but weaker in its recurrence than Thomas, reflecting the lower average associative centrality. Again, the number of unique ideas generated are only about 60% (39/57). The pattern for Fitzgerald is very different, with fewer generated ideas but a greater fraction of them unique (35/39). There is little tendency to search for extended periods within a region, though there are short bursts of local search. Finally, the recurrence pattern for Orwell is completely different, reflecting the chain-like structure of

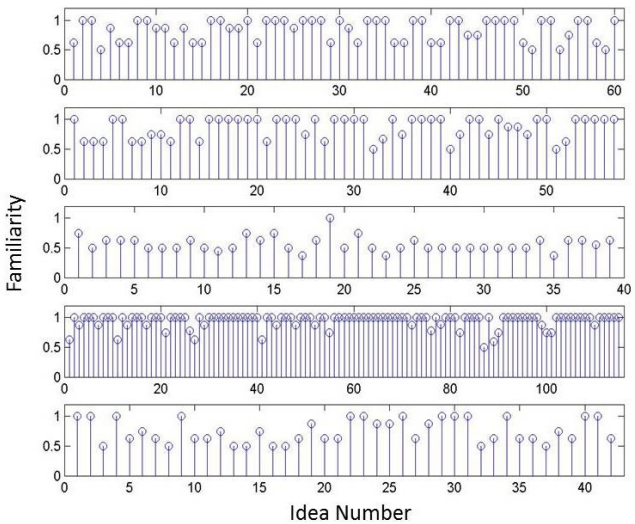

Fig. 9. Prior familiarity of ideas generated by the four authors. The results are for the simulations that generated the data in Figures 1 and 6. From top to bottom, the graphs are for: Thomas, Gay, Fitzgerald, Orwell and Thomas-Fitzgerald.

the network. A very large number of ideas is generated and most of them are unique (93/115). Essentially, the network systematically goes through all the modules representing snippets of the original text and generates them. Figure 7 shows the recurrence plots for four different runs of Thomas, showing the similar pattern of recurrence and local search. The dynamics for the other authors were similarly repeatable in a qualitative sense.

To highlight the effect of weight distribution (independent of connectivity) on the ideation dynamics, we did the following somewhat whimsical experiment. The weights in the Thomas network were replaced by those in the Fitzgerald network while maintaining the rank-ordering of weights for each word. Thus, if word j was the mth most frequent word in the Thomas corpus and the kth strongest association for j in the Thomas network was with word i, the weight from j to i in the Thomas network was replaced by that of the kth strongest associate in the Fitzgerald network for the mth most frequent word in the Fitzgerald corpus. Thus, while the association pattern from all words remained as for Thomas, the weight values became those of Fitzgerald. A typical recurrence plot for the resulting network is shown in the right panel of Figure 8, with a typical Thomas plot on the left for comparison. The hybrid system generated 42 ideas, of which 34 were unique, but the recurrence pattern of Thomas has been replaced by that of Fitzgerald, even though the connectivity is the same.

Finally, we look at the degree to which the ideas generated by the various networks are truly novel. The words in each idea are compared with all semantic units in the corresponding corpus and the best match found. The familiarity of the idea is measured as the fraction of its words that are covered in the best matching semantic unit. Figure 9 shows the results for the four simulations whose recurrence plots are shown in Figure 3, and for the Thomas-Fitzgerald simulation shown in Figure 8 (right panel). Not surprisingly, many ideas found were already present in the original corpus, but all networks generated some relatively unfamiliar ideas as well. The Fitzgerald network was the most likely to generate unfamiliar ideas and the Orwell network the least – as discussed above. The high proportion of familiar ideas generated by the Thomas and Gay networks reflect the broader definition of "semantic unit" in these compared to the prose authors, where the semantic units almost always corresponded to sentences. For example, a semantic unit in Thomas could be a long, unpunctuated stanza, where it would not be surprising to find all the words in an 8-word idea, but this is unlikely to happen in prose sentences. Surprisingly, however, the generative pattern of the Thomas network is altered drastically when the weight values from the Fitzgerald network are substituted. It should be remembered that the connectivity pattern and the *relative* weight order in this network is still that of Thomas; only the weight *values* have been changed. However, this causes the network to generate more unfamiliar ideas, showing that the ideas produced by the original Thomas network reflected its detailed associative structure and not just its gross connectivity.

VII. Conclusions

The main conclusions we draw from this very preliminary study are:

1. Associative networks built from the more 'creative' corpora, i.e., poetry, indeed show a more fat-tailed distribution of both degree and associative weights than corpora of prose. However, this observation needs to be corroborated through more and larger corpora.
2. The more fat-tailed poet networks lead to a complex, recurrent search for conceptual combinations, whereas the steeper prose weight distribution leads to a more linear search. It appears that the implicit correlations generated in the former case are more useful and meaningful than in the latter. Of course, detailed studies with larger corpora, more sophisticated network construction, and semantic analysis are needed to confirm this speculation.
3. The pattern of search depends not only on the pattern of association between concepts but also on the specific weights of these associations. The higher associative centrality of the poet networks seems to correlate with the cautious exploration shown by these networks. Decreasing this centrality by changing the weights changes the dynamics. Further studies will be needed to determine if – and how – associative weights reflect the style of thinking in an individual.

The present study can best be seen as a beginning, useful mainly for generating questions and highlighting the need for further study. Linking epistemic structure to thought and creativity remains a promising but elusive goal which, we hope, we will approach as our studies continue.

Acknowledgment

The authors would like to thank Laxmi Iyer, Simona Doboli, Paul Paulus, Alex Doboli, Dan Levine, Jared Ken-

worthy, Amer Ghanem, Marwa Shekfeh and Youyou Han for their contributions to this work and collboration in research of which it is a part. The authors also thank Yoed Kenett and Dror Kenett for their extremely useful suggestions. This paper is dedicated to the memory of Vince Brown, whose broad knowledge and deep insight were crucial to the genesis of this research.

REFERENCES

[1] D. T. Campbell, "Blind variation and selective retention in creative thought as in other knowledge processes," *Psychol. Rev.*, vol. 67, pp. 380–400, 1960.

[2] V. Brown, M. Tumeo, T. Larey, and P. Paulus, "Modeling cognitive interactions during group brainstorming," *Small Group Research*, vol. 29, pp. 495–526, 1998.

[3] D. K. Simonton, "Scientific creativity as constrained stochastic behavior: the integration of product, person, and process perspectives," *Psychol. Bull.*, vol. 129, pp. 475–494, 2003.

[4] G. Fauconnier and M. Turner, *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*. Basic Books, 2003.

[5] D. K. Simonton, "Creative thought as blind-variation and selective-retention: Combinatorial models of exceptional creativity," *Physics of Life Reviews*, vol. 7, pp. 156–179, 2010.

[6] S. Mednick, "The associative basis of the creative process," *Psychological Review*, vol. 69(3), pp. 220–232, 1962.

[7] V. Brown and P. B. Paulus, "A simple dynamic model of social factors in group brainstorming," *Small Group Research*, vol. 27, pp. 91–114, 1996.

[8] N. Marupaka and A. A. Minai, "Connectivity and creativity in semantic neural networks," in *Proceedings of IJCNN 2011*, 2011, pp. 3127–3133.

[9] N. Marupaka, L. R. Iyer, and A. A. Minai, "Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking," *Neural Networks*, vol. 32, pp. 147–158, 2012.

[10] A. E. Motter, A. P. S. de Moura, Y. C. Lai, and P. Dasgupta, "Topology of the conceptual network of language," *Physical Review E*, vol. 65, p. 065102(R), 2002.

[11] M. Sigman and G. A. Cecchi, "Global organization of the wordnet lexicon," *PNAS*, vol. 99, pp. 1742–1747, 2002.

[12] M. Steyvers and J. Tenenbaum, "The large scale structure of semantic networks: Statistical analyses and a model of semantic growth," *Cognitive Science*, vol. 29, pp. 41–78, 2005.

[13] M. E. Bales and S. B. Johnson, "Graph theoretic modeling of large-scale semantic networks," *Journal of Biomedical Informatics*, vol. 39, pp. 451–464, 2006.

[14] J. P. Guilford, "Creativity," *American Psychologist*, vol. 5, pp. 444–454, 1950.

[15] M. Boden, *The Creative Mind: Myths and Mechanisms*. Routledge, 2004.

[16] B. A. Nijstad and W. Stroebe, "How the group affects the mind: A cognitive model of idea generation in groups," *Personality and Social Psychology Review*, vol. 3, pp. 186–213, 2006.

[17] M. A. Schilling, "A small-world network model of cognitive insight," *Creativity Res. J.*, vol. 17, pp. 131–154, 2005.

[18] W. Duch, "Intuition, insight, imagination and creativity," *IEEE Comput. Intell.*, pp. 40–52, 2007.

[19] S. L. Thaler, "Neural nets that create and discover," *PC AI*, vol. May/June, pp. 16–21, 1996.

[20] L. R. Iyer, S. Doboli, A. A. Minai, V. R. Brown, D. S. Levine, and P. B. Paulus, "Neural dynamics of idea generation and the effects of priming," *Neural Networks*, vol. 22, pp. 674–686, 2009.

[21] P. Thagard and T. C. Stewart, "The aha! experience: Creativity through emergent binding in neural networks," *Cognitive Science*, vol. 35, pp. 1–33, 2011.

[22] M. Diehl and W. Stroebe, "Productivity loss in brainstorming groups: Toward the solution of a riddle," *Journal of Personality and Social Psychology*, vol. 53, pp. 497–509, 1987.

[23] M. D. Mumford and S. B. Gustafson, "Creativity syndrome: Integration, application, and innovation," *Psychological Bulletin*, vol. 103, pp. 27–43, 1988.

[24] M. I. Mobley, L. M. Doares, and M. D. Mumford, "Process analytic models of creative capacities: Evidence for the combination and reorganization process," *Creativity Research Journal*, vol. 5, pp. 125–155, 1992.

[25] P. B. Paulus and H. Yang, "Idea generation in groups: A basis for creativity in organizations," *Organizational Behavior and Human Decision Processes*, vol. 82, pp. 76–87, 2000.

[26] T. B. Ward, "Creative cognition, conceptual combination, and the creative writing of stephen r. donaldson," *American Psychologist*, vol. 56, pp. 350–354, 2001.

[27] V. Brown and P. Paulus, "Making group brainstorming more effective: Recommendations from an associative memory perspective," *Current Directions in Psychological Science*, vol. 11, pp. 208–212, 2002.

[28] K. L. Dugosh and P. B. Paulus, "Cognitive and social comparison processes in brainstorming," *Journal of Experimental Social Psychology*, vol. 41, pp. 313–320, 2005.

[29] M. Jung-Beeman, E. Bowden, J. Haberman, J. Frymiare, S. Arambel-Liu, R. Greenblatt, P. Reber, and J. Kounios, "Neural activity when people solve verbal problems with insight," *PLoS Biology*, vol. 2, pp. 0510–0510, 2004.

[30] E. Bowden, M. Jung-Beeman, J. Fleck, and J. Kounios, "New approaches to demystifying insight," *Trends in Cognitive Sciences*, vol. 9, pp. 322–328, 2005.

[31] D. O. Hebb, *Essay on Mind*. Lawrence Erlbaum, 1980.

[32] A. A. Minai, L. R. Iyer, D. Padur, and S. Doboli, "A dynamic connectionist model of idea generation," in *Proceedings of IJCNN 2009*, 2009, pp. 2109–2116.

[33] S. Doboli, V. R. Brown, and A. A. Minai, "A conceptual neural model of idea generation," in *Proceedings of IJCNN 2009*, 2009, pp. 723–729.

[34] L. R. Iyer, A. A. Minai, S. Doboli, V. R. Brown, and P. B. Paulus, "Effects of relevant and irrelevant primes on idea generation: A computational model," in *Proceedings of IJCNN 2009*, 2009, pp. 1380–1387.

[35] L. R. Iyer, V. Venkatesan, and A. A. Minai, "Neurcognitive spotlights:configuring domains for ideation," in *Proceedings of WCCI 2010*, 2010, pp. 3026–3033.

[36] J. P. Guilford, *The Nature of Human Intelligence*. McGraw-Hill, 1967.

[37] T. M. Amabile, *The Social Psychology of Creativity*. Springer-Verlag, 1983.

[38] D. K. Simonton, *Scientific Genius: A Psychology of Science*. Cambridge University Press, 1988.

[39] Y. Kenett, D. Kenett, E. Ben-Jacob, and M. Faust, "Global and local features of semantic networks: Evidence from the Hebrew mental lexicon," *PLoS ONE*, vol. 6, p. e23912, 2011.

[40] J. G. W. Raaijmakers and R. Shiffrin, "Search of associative memory," *Psychological review*, vol. 88, 1981.

[41] D. L. Nelson, V. M. McKinney, N. R. Gee, and G. A. Janczura, "Interpreting the influence of implicitly activated memories on recall and recognition," *Psychological Review*, vol. 105, pp. 299–324, 1998.

[42] A. Morais, H. Olsson, and L. Schooler, "Mapping the structure of semantic memory," *Cognitive Science*, vol. 2012, pp. 1–21, 2012.

[43] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statisical Mechanics*, vol. 2008, p. P10008, 2008.

[44] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, vol. 393, pp. 440–442, 1998.

[45] L. Li, D. Alderson, R. Tanaka, J. Doyle, and W. Willinger, "Towards a theory of scale-free graphs: Definition, properties, and implications (extended version)," California Institute of Technology, Tech. Rep. CIT-CDS-04-006, 2005.

[46] M. V. Tsodyks and H. Markram, "The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability," *Proceedings of the National Academy of Sciences USA*, vol. 94, pp. 719–723, 1997.

[47] R. S. Zucker and W. G. Regehr, "Short-term synaptic plasticity," *Annual Review of Physiology*, vol. 64, pp. 355–405, 2002.

[48] L. F. Abbott and W. G. Regehr, "Synaptic computation," *Nature*, vol. 431, pp. 796–803, 2004.

[49] J. Starzyk, "Mental saccades in the cognitive process," in *Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA*, 2011, pp. 495–502.

[50] R. Plotkin, *The Genie in the Machine: How Computer-Automated Inventing is Revolutionizing Law and Business*. Stanford University Press, 2009.