ARTICLE IN PRESS

Neural Networks ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2012 Special Issue

Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking

Nagendra Marupaka, Laxmi R. Iyer, Ali A. Minai *

School of Electronic and Computing Systems, University of Cincinnati, Cincinnati, OH 45221, United States

ARTICLE INFO

Keywords:
Semantic networks
Semantic cognition
Creativity
Cognitive dynamics
Itinerant dynamics
Attractor networks

ABSTRACT

Understanding cognition has been a central focus for psychologists, neuroscientists and philosophers for thousands of years, but many of its most fundamental processes remain very poorly understood. Chief among these is the process of thought itself: the spontaneous emergence of specific ideas within the stream of consciousness. It is widely accepted that ideas, both familiar and novel, arise from the combination of existing concepts. From this perspective, thought is an emergent attribute of memory, arising from the intrinsic dynamics of the neural substrate in which information is embedded. An important issue in any understanding of this process is the relationship between the emergence of conceptual combinations and the dynamics of the underlying neural networks.

Virtually all theories of ideation hypothesize that ideas arise during the thought process through association, each one triggering the next through some type of linkage, e.g., structural analogy, semantic similarity, polysemy, etc. In particular, it has been suggested that the creativity of ideation in individuals reflects the qualitative structure of conceptual associations in their minds. Interestingly, psycholinguistic studies have shown that semantic networks across many languages have a particular type of structure with small-world, scale free connectivity. So far, however, these related insights have not been brought together, in part because there has been no explicitly neural model for the dynamics of spontaneous thought. Recently, we have developed such a model. Though simplistic and abstract, this model attempts to capture the most basic aspects of the process hypothesized by theoretical models within a neurodynamical framework. It represents semantic memory as a recurrent semantic neural network with itinerant dynamics. Conceptual combinations arise through this dynamics as co-active groups of neural units, and either dissolve quickly or persist for a time as emergent metastable attractors and are recognized consciously as ideas. The work presented in this paper describes this model in detail, and uses it to systematically study the relationship between the structure of conceptual associations in the neural substrate and the ideas arising from this system's dynamics. In particular, we consider how the small-world and scale-free characteristics influence the effectiveness of the thought process under several metrics, and show that networks with both attributes indeed provide significant advantages in generating unique conceptual combinations.

1. Introduction

The "train of thought" or the "stream of consciousness" is an experience common to all humans – and probably to most other complex animals. Thoughts can be mundane or creative, transient or memorable, insignificant or salient – all emerging somehow from the continuous activity of billions of neurons in the brain influenced by the experience of the body embedded in its environment. How does this happen, and what determines the nature of the thoughts that arise in this way? Answering these questions is fundamental to any understanding of human cognition, and is the main motivation for the work reported in this paper.

Current understanding in neuroscience suggests that perception, thought and action are essentially the same phenomenon—a pattern of activity across complex networks of neural elements. When these elements are connected to the musculoskeletal system, the result is action. If this connection is (temporarily and voluntarily) disabled, one gets pure thought. This capability for "internal action" disconnected from overt behavior is the essential attribute that allows humans to think in the abstract, make plans, evaluate choices, generate ideas and solve complex problems. However, unlike behavior, thought is very difficult to study experimentally *as a temporal process*. Almost all studies of cognition rely on observables such as memory recall, response times, choice patterns, neural correlates of behavior, etc., but the thinking

* Corresponding author. Tel.: +1 513 556 4783.
E-mail addresses: minaiaa@gmail.com, Ali.Minai@uc.edu (A.A. Minai).

0893-6080/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2012.02.004

process itself remains mysterious and hard to observe (Graziano, Polosecki, Shalom, & Sigman, 2011). Computational modeling provides a principled way to address this issue (McClelland & Rogers, 2003; Tyler & Moss, 2001)—not so much to explain its details, but to provide general insight, much as attractor network models helped clarify the mechanisms of associative memory (Amit, 1989; Hopfield, 1982). In this spirit, the work presented in this paper describes a simple model for the thought process, informed by theoretical work on creative ideation (Brown, Tumeo, Larey, & Paulus, 1998; Campbell, 1960; Mednick, 1962; Nijstad & Stroebe, 2006; Schilling, 2005; Simonton, 2003, 2010) and recent empirical analysis of universal semantic structure in language (Bales & Johnson, 2006; Motter, de Moura, Lai, & Dasgupta, 2002; Sigman & Cecchi, 2002; Steyvers & Tenenbaum, 2005). In particular, it addresses two questions:

1. How do the conceptual associations embedded in the brain's semantic network generate a "train of thought", including creative thought?
2. How are the dynamical and functional characteristics of this process related to the structure of the semantic network?

In this paper, we assume amodal or lexical representations of concepts, and do not consider sensorimotor features. Research on the conceptual representations in the brain (Caramazza & Mahon, 2003; Damasio, 1989; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Kellenbach, Brett, & Patterson, 2001; Martin, 2007; Patterson, Nestor, & Rogers, 2007; Warrington & Shallice, 1984) indicates that modal, amodal and lexical representations exist in various parts of the cortex. These representations can be shown to arise from associative learning (e.g., in self-organized feature maps Kohonen, 1989). Higher level categorical representations can also arise from similar mechanisms (Iyer & Minai, 2011), and are explained partially by recent computational models (Ashby, 1998; Iyer & Minai, 2011; Kruschke, 1992; Love, Medin, & Gureckis, 2004; Nosofsky, Palmeri, & McKinley, 1994).

2. Theoretical formulation

The model we present is based on three basic assumptions:

1. All thought is homogeneous: This is the principle that creative and conventional thinking are the same *type* of phenomenon at the physiological level, differing not in their inherent phenomenology but only in the *value* of their output relative to utilitarian or aesthetic measures. This contrasts with the popular notion that genius uses mysterious processes unavailable to most individuals.
2. All thought is combinatorial: This principle asserts that all ideas are *conceptual combinations*, i.e., combinations of existing concepts or ideas. This is a central theme in most theories of creativity (Brown et al., 1998; Campbell, 1960; Fauconnier & Turner, 2003; Nijstad & Stroebe, 2006; Schilling, 2005; Simonton, 2003, 2010) and is discussed in more detail below.
3. All thought is associative: This is the principle that all new ideas arise b association with currently active ideas or percepts (Mednick, 1962), thus creating the "train of thought". The fundamentally associative nature of cognition is borne out by both experience and experiments, e.g., the body of research on priming (Collins & Loftus, 1975; Masson, 1995; McKoon & Ratcliff, 1992; Moss, Hare, Day, & Tyler, 1994; Plaut, 1995).

These assumptions may well be too simplistic, but we believe that they apply broadly enough to form the basis of a useful model. Given these assumptions, we argue that the difference between creative and mundane thinking arises not from the nature of the underlying process but from the organization of knowledge within the brain and, possibly, from differences in modulating factors such as inhibition, emotion, etc.

Studies of cognitive dynamics based on behavioral experiments indicate that it has an *itinerant* character (Tsuda, 2001), where periods of relative stability are interspersed with those of rapid transition. In particular, a recent study showed that a complex cognitive process can be segmented into saccade-like sequences of fixation and transition in the cognitive state (Graziano et al., 2011). These are similar to the idea of "mental saccades" proposed by Starzyk (Starzyk, 2011), and the underlying dynamics can be seen as itinerant traversal over a landscape of metastable states (Rabinovich et al., 2001). Over the last few years, as part of a larger cognitive model, we have developed an attractor network model for the itinerant emergence of conceptual combinations in a neural semantic network (Iyer et al., 2009; Iyer, Minai, Doboli, Brown, & Paulus, 2009; Iyer, Venkatesan, & Minai, 2010), and have recently presented a preliminary study of how the dynamics of this network is affected by its structure (Marupaka & Minai, 2011). The present paper builds on this work through more systematic experiments, and applies it to a real-world dataset.

3. Background

This section provides brief background on three important research areas that the work described in this paper draws upon.

3.1. Complex networks

Over the last few years, complex networks have become a major focus of study in many research areas, including physics, biology, sociology, linguistics, psychology, neuroscience, computer science and engineering (Barabasi, 2002; Newman, 2010; Newman, Strogatz, & Watts, 2001; Newman, Watts, & Strogatz, 2002; Watts, 1999, 2003). One interesting class of such networks are *scale-free networks*, characterized by a power-law distribution of node degree (number of connections for each node) (Albert & Barabasi, 2002; Barabasi & Albert, 1999), i.e., $p(k) \sim k^{\beta}$, where $p(k)$ is the probability of a node having k connections. *Random networks* – also known as Erdös–Renyi networks – with homogeneous uniform probability of connection between all node pairs have a Poisson degree distribution with a pronounced mode and an exponential tail, whereas the power-law distribution has a "fat tail". Scale-free networks are known to be generated by processes of preferential attachment, where new nodes entering the network connect preferentially to already well-connected nodes (Barabasi & Albert, 1999). This produces *hub nodes* with very high connectivity, which comprise the fat tail of the degree distribution. Many real-world networks, including the Internet and the World-Wide Web, have been found to be scale-free (Newman, 2010; Newman et al., 2001; Yook, Ad Jeong, & Barabasi, 2002).

Another widely studied class of networks is that of *small-world networks* (Watts, 1999; Watts & Strogatz, 1998), which are networks with a high *clustering coefficient*, C, but low *mean shortest path length* (MSPL), L, between node pairs. The MSPL is calculated by finding and averaging the shortest paths (in hops) between all pairs of nodes. The clustering coefficient is defined as the mean probability that the direct neighbors of two directly connected nodes are also directly connected. It is calculated as:

$$C = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i(n_i-1)}\sum_{j\in H_i} c_{ij} \quad (1)$$

where N is the number of nodes in the network, H_i is the set of nodes directly connected to node i, n_i is the size of this set, and $c_{ij} = 1$ if a connection exists from j to i, else 0. It should be noted that this calculation works for both undirected and directed networks. However, nodes in directed networks typically have distinct neighborhoods for incoming and outgoing connections, each giving its own clustering coefficient.

Please cite this article in press as: Marupaka, N., et al. Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. Neural Networks (2012), doi:10.1016/j.neunet.2012.02.004

Clustering is defined Random networks typically have low L and low C, whereas networks with only local connectivity have high L and C. Small-world networks, with low L and high C, thus represent a distinct class, and are known to be good models for many real-world complex networks such as power grids, neural networks and social networks (Newman, 2010; Watts & Strogatz, 1998). Small world characteristics have been found in cortical neural networks (Sporns & Zwi, 2004), and attractor neural networks with small world architecture are known to be very efficient as associative memories (Bohland & Minai, 2001). It has also been suggested that small-world connectivity in the brain may underlie high creativity in individuals (Schilling, 2005).

Scale-free networks have low L, but not necessarily low C, while small world networks do not necessarily have a power law degree distribution. However, it is possible to have networks that are both scale-free and small-world, and this structure has been found in many semantic networks (see below).

3.2. *The dynamics of thought*

A widely held view of creativity is that it arises from the combination of disparate concepts or ideas (Brown et al., 1998; Campbell, 1960; Fauconnier & Turner, 2003; Mednick, 1962; Nijstad & Stroebe, 2006; Schilling, 2005; Simonton, 2003, 2010). In particular, an influential model for creativity is the *blind variation and selective retention* (*BVSR*) *model*, first proposed by Campbell in 1960 (Campbell, 1960), and subsequently developed by Simonton (1988, 2003, 2010) and others. This theory postulates that potentially creative ideas arise as conceptual combinations through blind variation, and only those selected based on some evaluation criterion are retained.

The role of novel combinations in creativity has been noted by many great innovators. For example, Einstein is quoted as saying, "Taken from a psychological viewpoint … combinatory play seems to be the essential feature in productive thought—before there is any connection with logical construction in words or other kinds of signs which can be communicated to others" (quoted in Mednick (1962) and Simonton (2003)). Similarly, Poincare gives the following description of his creative process: "Ideas arose in crowds; I felt them collide until pairs interlocked, so to speak, making a stable combination. … the only combinations that have a chance of forming are those where at least one of the elements is one of those atoms freely chosen by our will. Now, it is evidently among these that is found what I called the good combination. … among the great numbers of combinations blindly formed almost all are without interest and without utility" (Poincaré, 1921) (quoted in Simonton (2003)). The last quote points to both the benefits and costs of blind combinatorial search—it is rich in novelty but is likely to be very poor in truly useful, or even sensible, ideas. Not only creative thinking but all thinking must contend with the trade-off between the need to produce sensible ideas and the need to generate new ones. This is simply the classic exploration vs. exploitation dilemma that underlies all complex adaptive processes, including evolution, learning, behavior and thought.

Since conceptual combinations arise in the minds of individuals, they must depend in some way on the way these minds organize information. Given the assumption that ideas arise associatively, the key issue is the structure of associations between mental representations of concepts. Such associations are the result of prior experience, and are encoded in the brain through the long-term modification of synapses between neurons involved in the representation of different concepts. Presumably, individuals whose minds encode unusual associations between concepts are likelier to find them when needed. This idea was formalized by Mednick (1962) as an *associative hierarchy*, relating the remoteness of associations in an individual's mind with the strength of these associations. He argued that the hierarchy is *steep* in non-creative individuals, who make very strong associations between commonly connected concepts but do not make unconventional associations. Such individuals are likely to think mainly in conventional terms. The creative individual, in contrast, has a *flatter* association hierarchy with fairly strong associations between concepts generally considered remote from each other. One way to see these two cases is in terms of the distribution of associations between concepts. In the non-creative individual, the distribution is likely to be strongly peaked at a typical level with an exponential tail, indicating that most concepts have approximately the same (small) number of associations. In contrast, creative individuals have a significant number of concepts with a large number of associations, creating a *fat-tailed distribution* such as the power-law.

An important consideration in the quality of thought is *fixation*—the inability to break out of conventional ideas, presumably because the strongest associations are between commonly linked concepts (e.g., "bread and butter", "bat and ball", etc.) Fixation can, thus, be seen as cognitive cliche. An interesting aspect of Mednick's hypothesis is that, due to their flatter associative hierarchies, creative individuals may not only generate more unusual conceptual combinations but also fewer conventional ones, which may explain the idiosyncrasy popularly associated with creative people. For the same reason, individuals with broad but relatively shallow expertise may be more creative than those with deep and narrow expertise (Schilling, 2005). Indeed, experiments have shown that the inclusion of a few unconventional thinkers can enhance the creativity of a whole group even if these thinkers are not especially knowledgeable (March, 1991; Nemeth, 1995).

3.3. *The organization of semantic knowledge*

The organization of semantic knowledge has been studied along two complementary tracks: (1) Through associative recall tests with human subjects; and (2) Through the structural analysis of lexical semantic networks.

Word association or cued recall tests (Nelson, Bennett, Gee, Schreiber, & McKinney, 1993; Nelson, McKinney, Gee, & Janczura, 1998; Nelson, Schreiber, & McEvoy, 1992; Nelson & Xu, 1995; Raaijmakers & Shiffrin, 1981; Ratcliff & McKoon, 1994) have produced a rich body of data on association norms, leading to theories of associative recall (e.g., McKoon & Ratcliff, 1992; Nelson et al., 1993, 1998, 1992; Ratcliff & McKoon, 1994). These experiments and theories provide a well-grounded understanding of the associative structure of semantic memory at the cognitive and behavioral level. This is complemented by studies of the structure of semantic networks obtained from dictionaries and thesauri by several investigators (Bales & Johnson, 2006; Motter et al., 2002; Sigman & Cecchi, 2002; Steyvers & Tenenbaum, 2005), all indicating that a number of lexical networks have consistent and complex structure. In particular, all these studies have shown that lexical semantic networks possess two interesting attributes: (1) A small-world architecture; and (2) A power-law degree distribution. Steyvers and Tenenbaum (2005) proposed a plausible model of semantic evolution that generates such networks, and have suggested that these properties might generalize across all associative semantic networks. It is, therefore, interesting to ask if the structure found by these studies simply reflects the formative processes of semantic networks, or might they also provide functional advantages to a cognitive agent.

4. Model description

In this paper, we describe a simple but neurally plausible computational model for associative search through the space of conceptual combinations in semantic neural networks, and use it to compare the characteristics of the resulting search in abstract networks with five different types of connectivity: Random (RA), Localized (LO); small-world (SW); scale-free (SF); and Steyvers–Tenenbaum (ST). We then use our model with an actual semantic network (WD) constructed using experimental word association norms (Maki, 2003, 2008; Nelson, McEvoy, & Schreiber, 1998).

It is important to emphasize that our work differs from most models of associative recall (e.g., Nelson et al., 1998, 1992) in two ways: (1) We consider the recall of conceptual *combinations* rather than single associates; and (2) We consider free recall (or spontaneous thinking) rather than cued recall. However, since conceptual combinations can themselves be seen as more complex concepts, the model can be considered a distributed form of associative chaining.

4.1. Network model

The core of the model is a semantic neural network of *N concept units*, each representing one concept (e.g., a word or an object). Each unit can be seen as a neural assembly. Since the goal in this paper is to study the dynamics of networks with various types of connectivity, we begin with networks of abstract concepts instead of actual words. We then also evaluate the model using empirically obtained word association data.

The connections between concept units represent *directed associations*, with w_{ij} denoting the connection from unit j to unit i. For simplicity, we assume that the connections are binary and symmetric, i.e., either two concepts are mutually associated with a weight of 1 or not associated, with a weight of 0 (this condition is relaxed for the real-world network studied later). The output of unit i at time t is $x_i(t)$, and the net input to a unit i and time t is given by:

$$x_i(t) = \sum_{j=1}^{N} w_{ij}(t)x_j(t) + \gamma_{\text{noise}}\xi_i(t) \tag{2}$$

where x_j are the outputs from units j, $\xi_i(t)$ is uniform white noise between 0 and 1, and γ_{noise}, is a fixed gain parameter.

The *state* of concept unit i at time t is given by:

$$y_j(t) = \alpha y_j(t-1) + (1-\alpha)x_i(t) \tag{3}$$

where α is an inertial parameter, typically set to a value slightly below 1. The network has *competitive activity*, and the K non-refractory units with the highest $y(t) > 0$ are allowed to fire at time t. In order to avoid arbitrary thresholding, units with excitation levels very close to the Kth most excited unit are also allowed to fire. Thus, the competitive firing is better considered "soft K-of-N".

The output of unit i is calculated as:

$$x_i(t) = \begin{cases} 1, & \text{if } y_i(t) \in k \text{ most excited units} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Unit activity and excitability are modulated by two other processes: (1) Refractoriness; and (2) synaptic modulation.

Refractoriness captures the fact that periods of high activity for neurons deplete resources and must be followed by intervals of refractoriness to replenish these resources. In the model, once unit i fires, it may remain active for an *activity duration* φ if $y_i^c(t)$ remains sufficiently high (i.e., it keeps winning the competition), after which it enters a *refractory period*. This is modeled through a *resource*, $r_i(t) \in R(t)$, with the following dynamics:

$$r_i(t) = \begin{cases} (1-\lambda_-)r_i(t-1), & \text{if active} \\ r_i(t-1) + \lambda_+(1 - r_i(t-1)), & \text{if inactive} \end{cases} \tag{5}$$

where λ_- is the *resource depletion rate*, and λ_+ is the *resource recovery rate*. A neuron is said to be in a refractory state if:

$$r_i(t) \le \Theta_r. \tag{6}$$

Thus, a unit's resource is depleted when it fires and recovers when it is inactive due to lack of stimulus or refractoriness.

Synaptic modulation, i.e., activity-dependent short-term change in synaptic strength, has recently been proposed as an important component of neural information processing and short-term memory (Abbott & Regehr, 2004; Zucker & Regehr, 2002). In the model, synapses that are excited repeatedly by pre-synaptic activity temporarily become habituated and weaken while the activity persists, recovering gradually when activity ceases. This is modeled as follows:

$$w_{ij}(t) = \begin{cases} (1-\psi_-)w_{ij}(t-1), & \text{if active} \\ w_{ij}(t-1) + \psi_+\left[w_{ij0} - w_{ij}\right], & \text{if inactive} \end{cases} \tag{7}$$

where ψ_- and ψ_+ represent the synaptic decay and recovery rates respectively and w_{ij0} represents the initial weight of the synapse.

4.2. Network dynamics

The dynamics of the network emerges from an interplay between the recurrent excitation within the network, the competitive K-of-N activity rule, and the modulatory processes (refractoriness and synaptic modulation). If the currently active group of (approximately) K units form a sufficiently strongly connected set, the group can sustain itself under competition, and only becomes inactive eventually due to resource depletion and synaptic modulation. This can be seen as an *emergent metastable attractor*. In contrast, if a group of relatively weakly connected units become co-active, they cannot remain active as a group and quickly fall apart. Thus, the dynamics of the network is *itinerant* (Tsuda, 2001) or "sticky", with periods of metastable activity patterns punctuated by intervals of transient activity.

Since ideas are defined as conceptual combinations and each unit in the network represents a concept, any set of co-active units can be seen as a potential idea. However, we assume that only those co-active sets that persist beyond a certain duration, Θ_{th} – termed the *awareness threshold* – are perceived consciously as ideas, whereas the rest remain subconscious. The intuition is that the persistence of an active concept group as a metastable attractor indicates that these concepts have sufficiently strong mutual associations based on prior experience and form a coherent idea, while transient groups comprise poorly associated concepts. It should be noted that, in contrast with standard associative memory networks (Hopfield, 1982), no attractors are explicitly embedded in our network. They are configured *latently* by the accumulation of associations between individual concepts. Over time, the network comes to embed a large number of such *latent ideas*, which are unmasked emergently by the dynamics of the system. Thus, in a sense, any novel ideas that emerge during thinking are "false memories"—ideas that were never explicitly experienced but are constructed from fragments of previous ideas that happen to connect together. Such combinatorial spurious memories are well-known in attractor networks (Amit, 1989), and are generally regarded as a nuisance when the goal is to store and recall specific memories. We assert that, for a system to be capable of creativity, such spurious memories are, in fact, very valuable, and their exclusion – e.g., by repeated reinforcement of "true" memories at their expense – reduces creativity. Creative thinking is productive confabulation. This point has also been made independently by others (Plotkin, 2009; Thaler, 1996a, 1996b).

Functionally, the dynamics of the system can be seen as generating a sequence of ideas with intervening periods of transience. We characterize these functional dynamics in terms of

Please cite this article in press as: Marupaka, N., et al. Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. Neural Networks (2012), doi:10.1016/j.neunet.2012.02.004

the following attributes:

1. *Productivity*, ρ, is the number of unique ideas generated over a finite period.
2. *Efficiency*, η, is a measure of how little time is "wasted" generating repeated ideas. This is calculated as:

   $$\eta = \rho/\rho_{\text{all}} \quad (8)$$

   where ρ_{all} is the total number of ideas (including repeated ones) generated by the network during the simulation. Low efficiency can be seen as indicating fixation.
3. *Consistency*, ξ, quantifies repeatability in the behavior of a network over distinct simulations. It measures the overlap between ideas generated over repeated runs from different initial conditions. A value closer to 1 indicates high repeatability, implying that the dynamics at the idea level is more "ergodic".
4. *Coherence*, ω, measures how strongly connected a generated idea is relative to the knowledge embedded in the network. It is measured by calculating the clustering – or mutual connectivity – among the nodes participating in the idea. If idea I_k has K active nodes,

   $$\omega(I_k) = q(I_k)/K(K-1) \quad (9)$$

   where $q(I_k)$ is the number of connections that exist between the K nodes active in I_k. This measure is modified slightly for the WD network, where: (1) All non-zero weights are set to 1 before calculation of coherence; and (2) The raw coherence value obtained is scaled by the connectivity of the WD network relative to the abstract networks. This allows the coherence for all networks to be compared.

Using cooperatively co-active groups of concepts as representations of ideas has been considered by Nelson et al. (1993) in the context of implicit memory. They found that concepts that are part of such highly connected (or clustered) groups are easier to recall than those that are not.

From a neurophysiological viewpoint, a more plausible instantiation of our co-activity model might be in terms of emergent synchronization among neuronal assemblies, which has been suggested as the brain's main mechanism for representational binding (Bressler & Tognoli, 2006; Eckhorn et al., 1988; Engel, Fries, & Singer, 2001; Singer et al., 1997; Singer & Gray, 1995; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Several computational models have been developed for such systems (e.g., Wang & Terman, 1995), but we use the simpler model described above for clarity.

Finally, it should be noted that, while we use the term "concept" for the information represented by each network unit, they could also be seen as "features". The distinction between features, concepts and ideas is largely a matter of the level in a representational hierarchy rather than an essential difference.

4.3. Network generation

The study of abstract networks centers on the Steyvers–Tenenbaum model, which represents the structure of actual lexical networks. Thus, we first generate this network for a specific value of N and other parameters, and then generate the other four architectures to be comparable to it. The networks are generated as follows:

Steyvers–Tenenbaum (ST) network: This network is generated using the algorithm described by Steyvers and Tenenbaum in Steyvers and Tenenbaum (2005). The process begins with M fully interconnected nodes, with the remaining $N - M$ nodes added subsequently one at a time. The addition of each new node requires two steps: (1) An existing node, i, is chosen with probability $P_i = k_i/\sum_j k_j$, where k_j is the degree of node j; (2) The new node is connected randomly to M other nodes from the set of nodes to which i is already connected. Thus, the new node partially replicates the connectivity of node i, which is chosen with a preference for higher degree nodes. The resulting network is both small-world and scale-free, with an exponent near 3. For the simulations described here, we use $N = 500$ and $M = 6$, giving a total of $n_c = 2979$ bidirectional connections for an extremely sparse network with mean node degree 0.023.

Random (RA) network: The RA network has N nodes with n_c connections between node pairs chosen randomly with equal probability. This results in a Poisson degree distribution for the nodes (Newman, 2010).

Localized (LO) network: For the LO network, nodes are first placed randomly in 2-dimensional Euclidean space in order to create distance relationships, each node making connections within a certain radius, r, of itself. The radius is chosen so that the total number of connections in the network is very close to n_c.

Small-World (SW) network: The SW network is obtained by randomly rewiring short connections in the LO network to more distant nodes until the mean clustering coefficient of the network equals that of the ST network.

Scale-Free (SF) network: The SF network is generated through a preferential attachment process (Barabasi & Albert, 1999), constrained by the requirement that the total number of connections be close to n_c. This results in a scaling exponent of approximately 3, albeit over a large range of node degrees, with a cutoff at very high values.

To summarize, all nodes have the same number of nodes and connections, the ST and SW networks have the same clustering coefficient, the LO and SW networks have the same basic connection radius, and the ST and SF networks have the same degree scaling exponent, albeit over a somewhat narrower range for the SF network.

The *WD network* is generated using actual word association measures generated by Maki (2003, 2008) based on the University of South Florida free association norms dataset (Nelson et al., 1998). The available dataset had word associations and word frequencies for 4374 nouns, of which we chose 2059 by excluding low frequency words. In the network, each word is represented by a node, and the connection weight, w_{ij}, from node j to node i equals the relative strength of their association. Thus, the weights are neither binary nor symmetric in this case. Also, the resulting network is considerably sparser than the abstract networks described above, with a mean node degree of 0.005. It should be noted that a larger network based on the same dataset was also analyzed by Steyvers and Tenenbaum (2005), but only with respect to structural properties.

Fig. 1 shows the degree distributions for all six networks.

4.4. Network attributes

Fig. 1(a)–(e) show the degree distributions for the five abstract networks, and Fig. 1(f) the in- and out-degree distributions for the WD network, all in log–log coordinates. It is clear that the RA, LO and SW networks have Poisson distributions with exponential tails while the ST and SF networks' distribution have (truncated) power law tails. It is interesting to interpret these connectivity patterns from the perspective of Mednick's associative hierarchies. In the three networks with Poisson degree distributions, all concepts have approximately similar numbers of associations, whereas the ST and SF networks have concepts with very different numbers of associations, including some with a very large set. These high degree "hub" nodes can serve to link disparate concepts.

Table 1 summarizes the structural properties of all the networks. As expected, the RA network has a low mean shortest path length (MSPL) and almost no clustering, whereas the LO

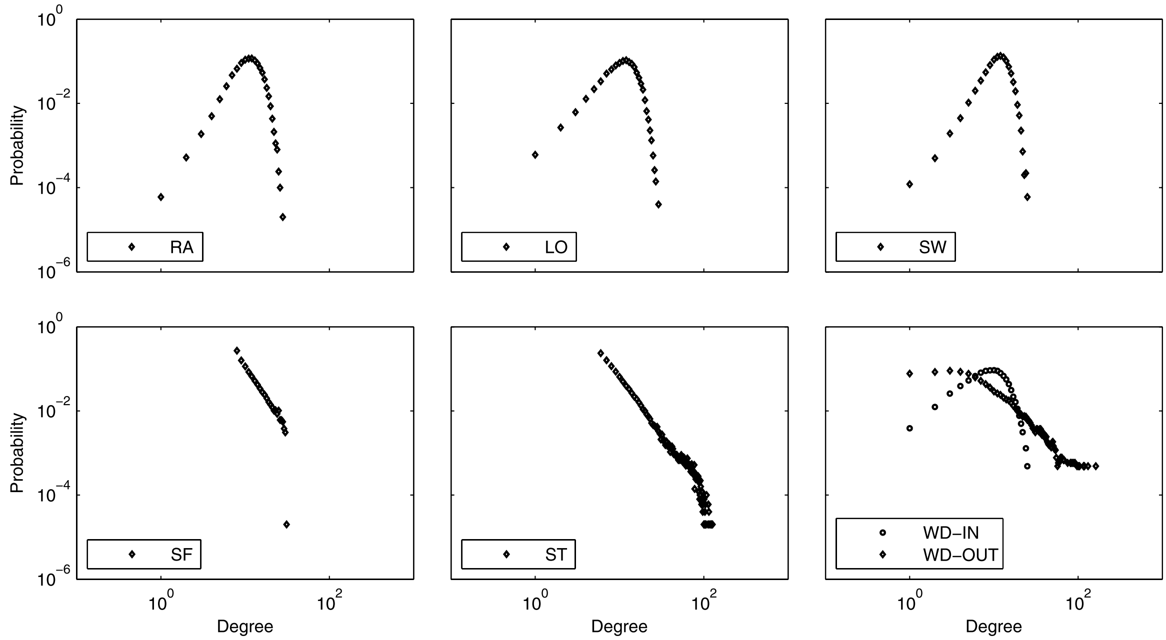

Fig. 1. Degree distribution for the six networks. Note that the WD network has directed links, resulting in separate distributions for in-degree and out-degree.

Table 1
Network properties.

Network	Size (N)	Mean degree (λ)	Clustering coefficient (C)	MSPL (L)	Radius (R)	Diameter ($\hat{D}$)	Scaling exponent (γ)
RA	500	12	0.0238	2.77	3.49	4.56	NA
LO	500	12	0.6160	7.78	9.56	19.6	NA
SW	500	12	0.2790	3	3.49	5.09	NA
SF	500	12	0.0290	2.78	3	4.01	3.2
ST	500	12	0.2780	2.79	3	5	2.5
WD	2059	10	0.1258	4.16	5.5	11	1.3 (in-degree)
			0.1464 (in)				
			0.1053 (out)				

network has very long paths and very high clustering. The ST and SW networks have MSPL similar to the RA networks but much higher clustering, indicating small-world characteristics (Watts & Strogatz, 1998). The SF network has low MSPL and very low clustering coefficient. Finally, the WD network also shows a small world pattern, though with a lower clustering coefficient and longer MSPL than the SW and ST networks. This is because, in contrast to the other networks, the connections in the WD network are directed. The values we find are in agreement with those found for a larger version of the WD network in Steyvers and Tenenbaum (2005).

5. Simulations and results

All five abstract network models were simulated using $N = 500$ units with connectivity generated as described above. The WD model with $N = 2059$ was also simulated with the same parameter values. There are three significant differences between the five abstract networks and the WD network:

1. The WD network is larger but also much sparser than the abstract networks. On average, each WD node receives 10 inward connections compared to 12 for each abstract network.
2. The WD network has directed, asymmetric connections.
3. The WD network has real-valued rather than binary weights, with the values indicating the strength of association between concepts.

In analyzing the simulations, it is important to remember that almost all the parameters for the WD network – unlike those for the abstract networks – are constrained by the data, and not every performance metric for this network can be compared directly with those for the abstract networks.

An issue in the simulations is to set the awareness threshold, Θ_{th}, appropriately given the other dynamical parameters in the system. To do this, we looked at the distribution of persistent activity duration for individual units with all parameter values fixed. This distribution was found to be bimodal as shown in Fig. 2 (see also Marupaka & Minai, 2011). In most instances, units stay active only for a few consecutive steps before switching off, which indicates transient activity. However, in some cases, the persistent activity lasts for durations between 65 and 80 consecutive time steps, which is the limit set by the resource decay rate. We hypothesize that these units are participating in a metastable attractor – i.e., an idea – by remaining active as long as physically possible. Based on this figure, we set the awareness threshold near the lowest point of the distribution at $\Theta_{th} = 40$.

Fig. 3 shows the mean values of the four metrics – productivity (η), efficiency (ρ), coherence (ω) and consistency (ξ) – for all six networks. Each data point is averaged over 20 independent runs. Fig. 4 shows the *throughput* for each system, defined as the fraction of time the system spends in metastable attractors that correspond to ideas. The distribution of coherence values over all ideas and unique ideas for each of the networks is plotted in Fig. 5. Fig. 6 shows two representative ideas from the WD network.

To illustrate the dynamics of the thought process as it moves from one idea to the next, Fig. 7 shows the activity of part of the WD network over the entire period spanning the two successive ideas, while Fig. 8 shows the Hamming distances of this activity from the first idea (top) and the second idea (middle), and between

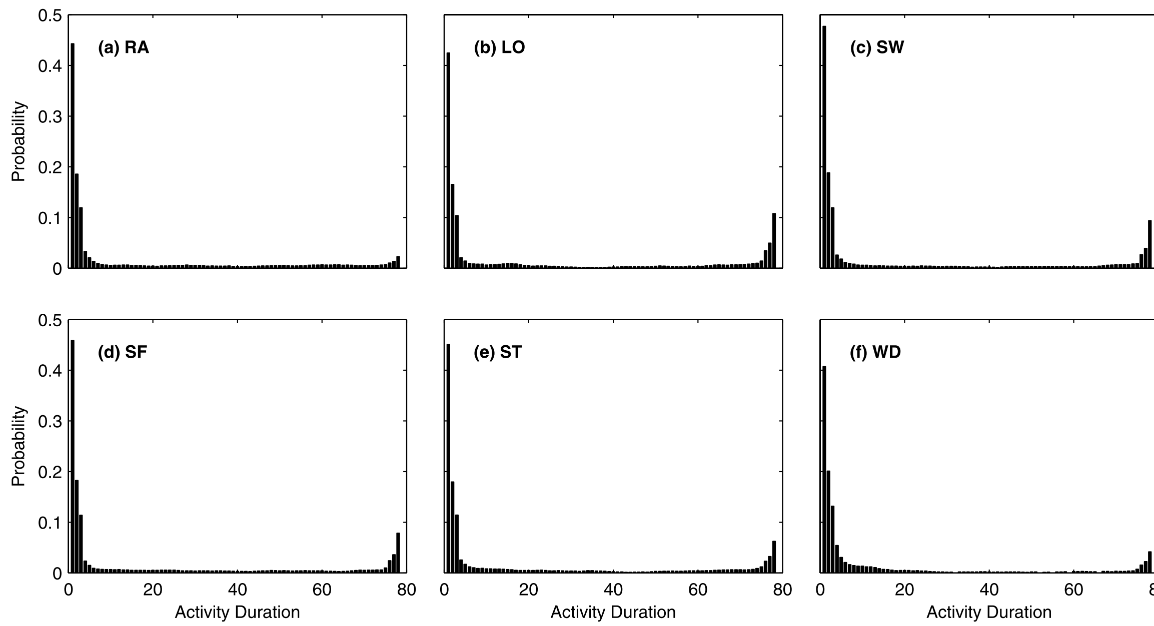

Fig. 2. Distribution of sustained activity periods for neurons in each of the six networks (for a single run). In each case, the distribution is strongly bimodal. The high peak on the left represents brief transient activations and the one on the right corresponds to sustained activations—usually indicating participation in a metastable attractor (idea).

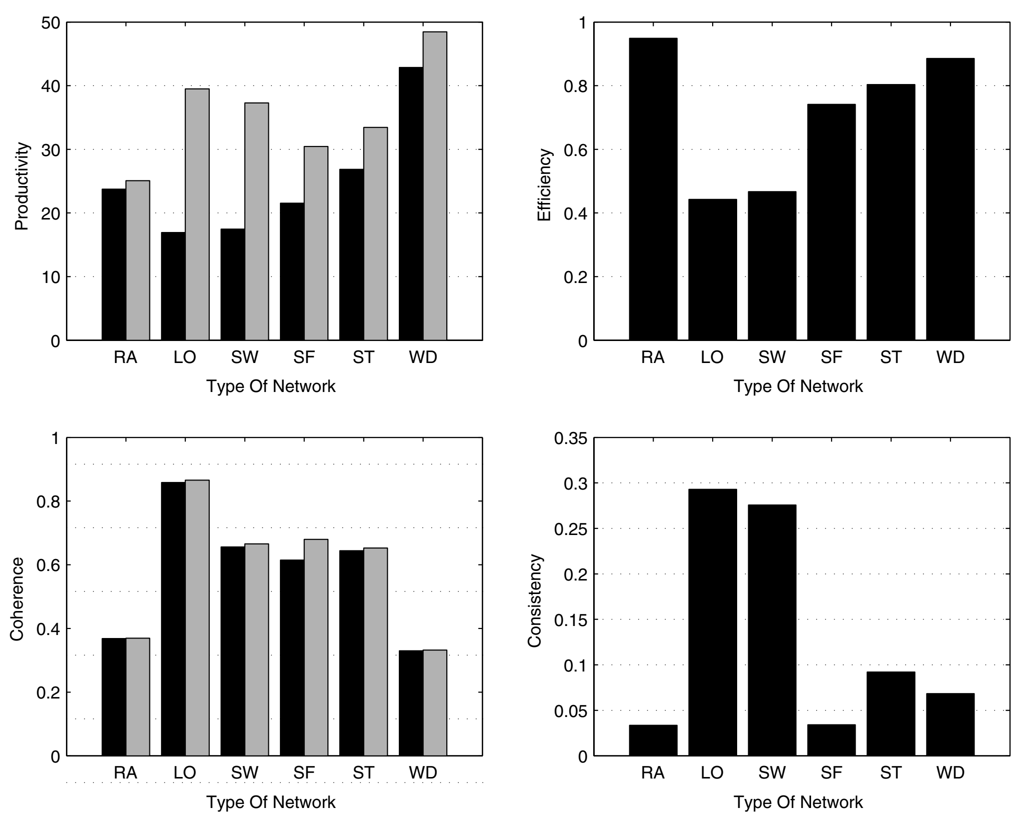

Fig. 3. The mean Productivity, Efficiency, Coherence and Consistency metrics for all six networks. Each network's data is averaged over 20 runs. In the plots on the left, solid gray bars correspond to "All" ideas and black bars to "Unique" ideas only.

successive states. The left side of both figures shows how an existing idea dissipates, and the right side shows the emergence of the next idea. Table 2 gives a more detailed view of the transition out of the first idea and into the second idea, showing the transient conceptual combinations generated as the first idea dissipates and those generated as the second idea emerges. As indicated by Fig. 8 and Table 2, neither happens abruptly, but requires only a short duration of dissipation or build-up. This is typical for all ideas seen in the simulation, though we have not explored this issue systematically. Looking at the dependence of the emergence and dissipation times on network characteristics is a potentially interesting issue for future research.

Table 2
Steps of conceptual combinations spanning two successive ideas (in bold).

T	N_0	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}
…	…	…	…	…	…	…	…	…	…	…	…
240	England	English	France	French	Guide	Language	Latin	London	Paris	Spanish	Tour
241	England	English	France	French	Guide	Language	Latin	London	Paris	Spanish	
242	**England**	**English**	**France**	**French**	**Greek**	**Language**	**Latin**	**London**	**Paris**	**Spanish**	
…	…	…	…	…	…	…	…	…	…	…	…
296	**England**	**English**	**France**	**French**	**Greek**	**Language**	**Latin**	**London**	**Paris**	**Spanish**	
297	England	English	France	French	Greek	Language	Latin	Literature	Paris	Spanish	
298	English	Europe	France	French	Greek	Language	Latin	Literature	Paris	Spanish	
299	English	Europe	France	French	Greek	Language	Latin	Literature	Paris	Spanish	
300	English	France	French	Greek	Language	Latin	Literature	Paris	Poetry	Spanish	
301	English	France	French	Greek	Language	Latin	Literature	Paris	Poetry	Spanish	
302	English	France	French	Greek	Language	Latin	Literature	Paris	Poem	Spanish	
303	English	French	Greek	Language	Latin	Paris	Poet	Poetry	Spanish	Verse	
304	English	French	Greek	Language	Latin	Literature	Poem	Poet	Spanish	Verse	
305	English	French	Greek	Language	Latin	Poem	Poet	Poetry	Spanish	Verse	
…	…	…	…	…	…	…	…	…	…	…	…
578	Cover	Discover	Find	Glance	Hide	Look	Search	See	Seek	Sight	
579	Cover	Discover	Discovery	Find	Glance	Hide	Lid	Look	Search	Seek	
580	Cover	Discover	Discovery	Find	Glance	Hide	Lid	Lose	Search	Seek	
581	Cover	Discover	Discovery	Find	Hide	Lid	Lose	Search	Seek	Win	
582	Defeat	Discover	Discovery	Find	Hide	Lid	Lose	Search	Seek	Win	
583	Defeat	Discover	Discovery	Find	Hide	Lose	Search	Seek	Victory	Win	
584	Defeat	Discover	Find	Gain	Hide	Lose	Search	Seek	Victory	Win	
585	Defeat	Discover	Find	Gain	Hide	Lose	Search	Seek	Victory	Win	
586	**Defeat**	**Discover**	**Find**	**Hide**	**Lose**	**Search**	**Seek**	**Triumph**	**Victory**	**Win**	
…	…	…	…	…	…	…	…	…	…	…	…

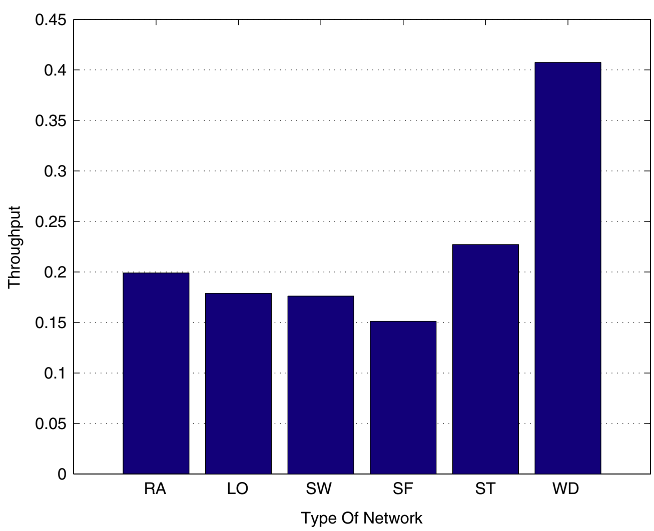

Fig. 4. The average throughput for the six networks. Throughput represents the fraction of time a network spends in a metastable state corresponding to an idea.

6. Discussion

Several interesting observations emerge from the results presented in the previous section:

Productivity and efficiency: Fig. 3(a) indicates that, while the LO and SW networks generate the greatest number of ideas among abstract networks, a large fraction of these are repeated. The RA network, in contrast, generates fewer ideas but almost all are unique. The SF network generates more unique ideas than LO or SW, though fewer than RA. The ST network, however, generates more unique ideas than RA, and the WD network generates even more. Comparing the efficiency of these networks, as shown in Fig. 3(b), it is clear that the RA network is most efficient, and the three other networks with power law degree distribution are also much more efficient than the two other networks with Poisson degree distribution (LO and SW).

Mean coherence: Fig. 3(c) shows that, on the average, the ideas produced by the LO, SW, SF and ST networks were much more internally coherent than those generated by the RA networks. The mean coherence for ideas from the WD network is also low, but this is mainly because this network, unlike the others, had directed links and much lower mean weights. Thus, the coherence values for the WD network should not be compared directly with the other networks. It is interesting to note that the mean coherence values for all ideas and unique ideas are virtually identical for all networks except the SF network. This suggests that the SF network tends to repeat a few highly coherent ideas disproportionately (see below).

Consistency: Fig. 3(d) indicates that the RA and SF networks are extremely inconsistent, producing very different ideas on each independent trial with the same connectivity pattern. Further investigation shows that the underlying reasons are quite different in the two cases. The RA network simply has a lot of random, weakly connected sub-networks that are not associated strongly with each other, making each trajectory through the network quite different. In contrast, the SF network has a few very strongly connected components, and the system gets trapped around one of these on each run. The networks with high clustering tend to have higher consistency, but this is weakened by power law connectivity.

Throughput: Fig. 4 shows that all networks have roughly similar throughput, spending between 14% and 22% of their time in metastable states. The WD network seems to have much higher throughput, possibly indicating a greater richness of ideas.

Coherence distribution: Fig. 5 shows that the distribution of coherence across ideas differs considerably in the six networks. As expected, coherence values for the RA network are tightly clustered around a low mean. The LO network, in contrast, preferentially produces highly coherent ideas, reflecting its high degree of clustering. It is very interesting to note that the introduction of shortcuts in modifying the LO architecture to the SW one completely changes the distribution of coherence, with the distribution taking on a Gaussian shape around a fairly high mean. Almost exactly the same distribution is seen in the ST network, indicating that it is characteristic of small-world networks. The distribution seen for SF networks is the most intriguing. The distribution for all ideas is strongly bimodal, with a bell-shaped lobe at low coherence values and a sharp peak towards high values. However, the distribution for unique ideas shows that most of the

Please cite this article in press as: Marupaka, N., et al. Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. Neural Networks (2012), doi:10.1016/j.neunet.2012.02.004

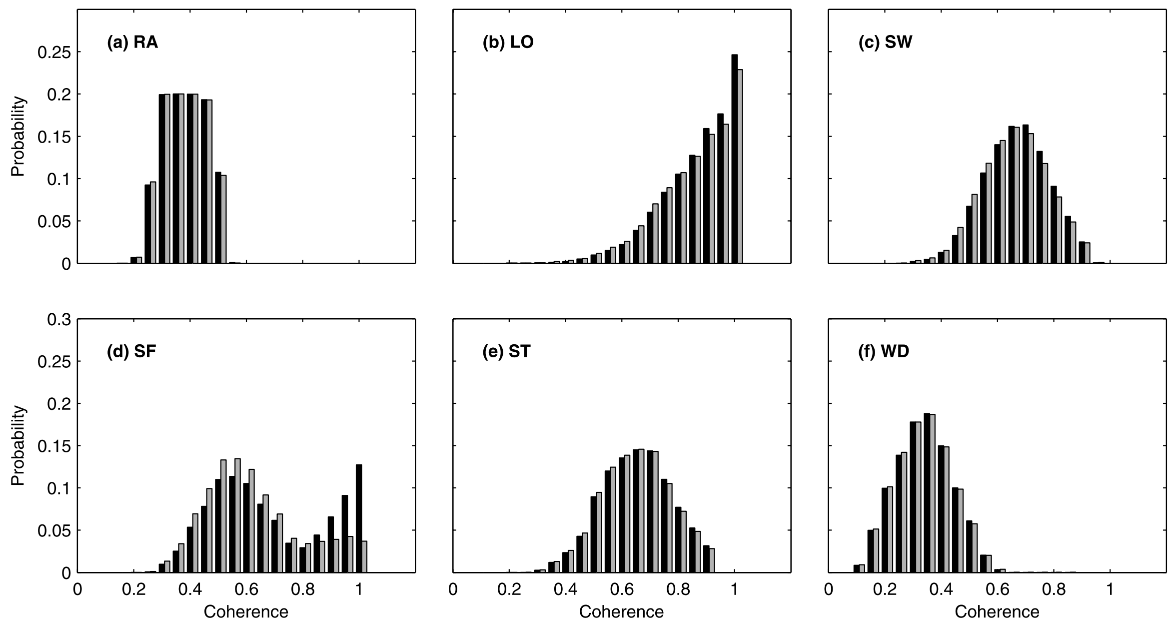

Fig. 5. The distribution of coherence across the ideas generated. Solid black bars correspond to "All" ideas and gray bars to "Unique" ideas only. See text for discussion.

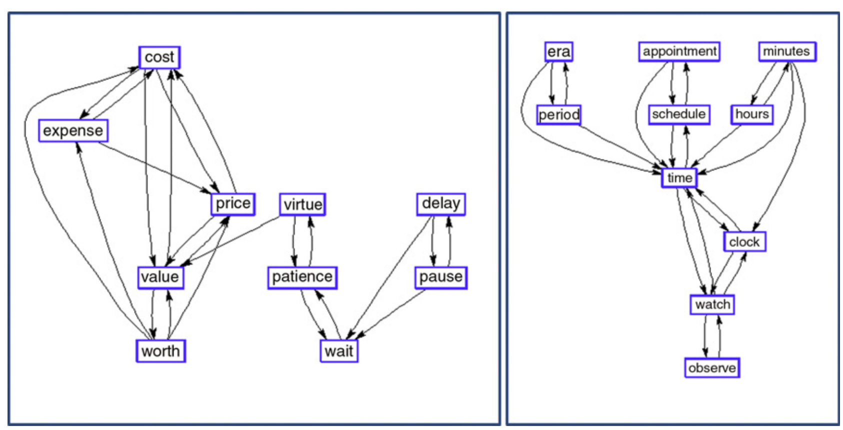

Fig. 6. Two sample ideas from the WD network.

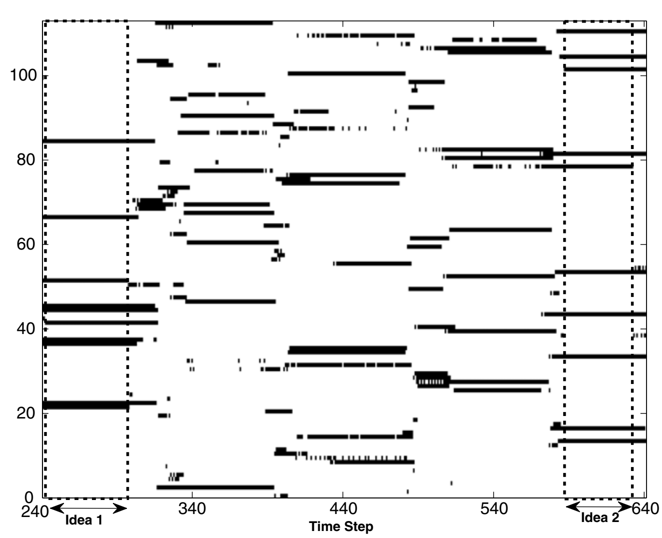

Fig. 7. The activity of 110 neurons from the WD network over the period spanning two successive ideas. The first idea occurs at the left extreme of the plot and the second at the right extreme, as indicated by the dotted bounding boxes. The activity in between represents transient conceptual combinations.

very high coherence ideas are just repetitions of the same few cases. Finally, as discussed earlier, the exact values of coherence for the WD network are not directly comparable with those in the five abstract networks, but the Gaussian shape of the distribution is suggestive of the ST network. This similarity is also borne out by the high efficiency of the WD network.

Ideas generated in the WD network: Since the WD network is based on word associations, the "ideas" generated from it tend to be combinations of semantically and functionally related words. More useful ideas would require a network with multiple types of connections between concepts/words and possibly some schematic mechanism. It is interesting to note, however, that even in this simplified setting, the presence of polysemic or multi-domain terms like "value" and "time" tend to connect concepts that might not otherwise be connected. This is a concrete example of how concepts in disparate ideas might come together through the agency of bridging concepts with multiple meanings.

We also looked at the distribution of durations for individual idea (i.e., the time for which a metastable pattern persisted) and the distribution of intervals between ideas. The latter were remarkably close to exponential, indicating a Markov-type process at the level of idea–idea transitions. The idea duration distributions were generally flat and did not show any interesting regularities.

Please cite this article in press as: Marupaka, N., et al. Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. Neural Networks (2012), doi:10.1016/j.neunet.2012.02.004

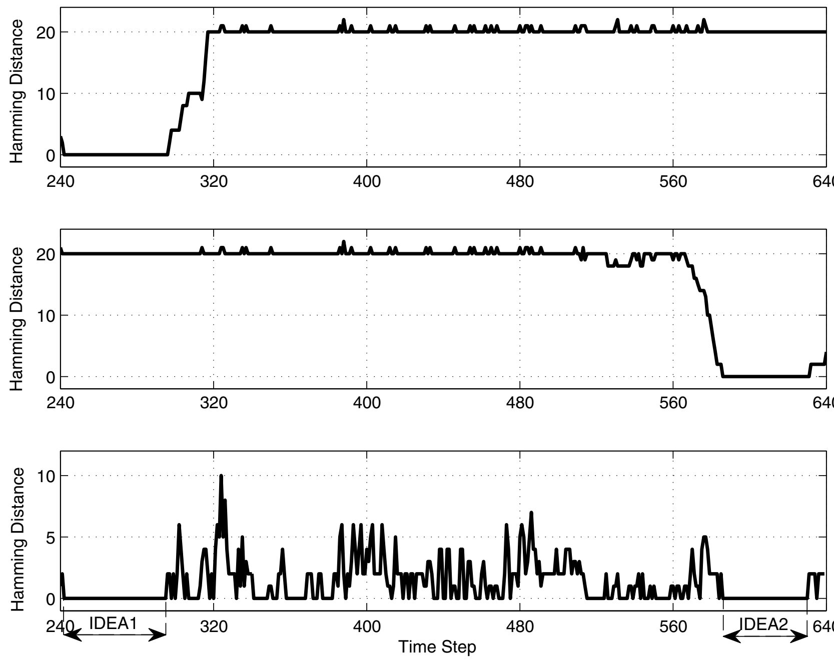

Fig. 8. Top. Hamming distance of current activity state from IDEA 1; Center. Hamming distance of current activity state from IDEA 2; Bottom. Hamming distance of each state from the previous state. Note that ideas do not dissipate or emerge abruptly, but do so over a short duration.

They are governed mainly by the values used for the modulation parameters.

All these observations lead to the following conclusions:

1. High clustering promotes fixation, i.e., repetition of the same ideas, but this can be mitigated by power law connectivity. Thus, the Steyvers–Tenenbaum architecture achieves high efficiency in spite of a high clustering coefficient.
2. High clustering leads to high mean coherence, indicating that highly clustered networks tend to produce "sensible" ideas.
3. Power law connectivity in the absence of high clustering (SF network) results in the generation of two classes of ideas: (a) Low coherence "novel" or "nonsensical" ideas similar to those generated by low clustering random networks; and (b) A small number of highly coherent, often repeated "cliche" ideas representing densely connected neighborhoods of hub nodes. Thus, a power law network of associations without high clustering is not very useful.
4. High clustering leads to consistent trajectories on different runs, but this effect is weakened significantly by power law connectivity. Non-power law connectivity without high clustering is not enough to produce high consistency.

Thus, it appears that networks with connectivity that is both small world and power law provide the best search performance of all the architectures considered. They combine the benefits of each attribute and mitigate the problems of each – small-world connectivity ensuring high coherence and power law ensuring high efficiency – leading to a very productive, efficient and consistent search for coherent ideas.

Since coherence is, by definition, a measure of consistency with prior knowledge, an idea with very high coherence represents an "old" idea. In contrast, an idea with very low coherence can be seen as bordering on absurdity—a quirk of the system's dynamics. However, ideas with a moderate degree of coherence can be regarded as sensible but novel. The ability to generate ideas in the mid-to-high range of coherence can, therefore, been seen as a signature of a healthy, creative mind, whereas a tendency towards low coherence (as in the RA network) indicates confusion, and a tendency towards very high coherence (as in the LO network) fixation. The pattern shown by the SF network is indicative of a mixture of confusion and fixation—a pattern characteristic of certain mental pathologies and dementias. This also suggests the very intriguing possibility that a healthy ST-type network with high clustering and power law connectivity could gradually become more like an SF network with the loss of some local connectivity, leading to a radical transformation in the balance between confusion, creativity, conventional thinking and fixation. This may have relevance for the understanding of various pathologies related to mental illness and dementia.

In summary, the following broad characterizations can be made for the type of "thinking" supported by each of the five abstract networks:

1. *Random*: Creative but nonsensical.
2. *Localized*: Inefficient and conventional to the point of obsessive fixation.
3. *Small-world*: Inefficient but spanning a healthy range between creativity and conventionality.
4. *Scale-free*: Mixture of creative-nonsensical and conventional-fixated.
5. *Steyvers–Tenenbaum (small-world and scale-free)*: Efficient and spanning a healthy range between creativity and conventionality.

7. Conclusion

In this paper, we have investigated the effect of connectivity on the dynamics of a neural model for generating conceptual combinations. The most significant result to emerge from the study was that networks with small-world power-law connectivity provide the best balance between the efficient search and the need to generate both novel and conventional ideas. Such connectivity has, in fact, been found in the study of several real-world semantic networks. Other types of connectivity were found to generate dynamics that may correspond to pathological states of mind. The semantic neural network model used in this study represents a promising way to model the dynamics of spontaneous thought. Because of its associative dynamics, the model is especially suitable for investigating the phenomenon of priming (Collins & Loftus, 1975), where external cues can strongly influence subsequent perceptions, thoughts, inferences, decisions. Spreading activation and attractor dynamics models have been applied to the study of priming (Masson, 1995; McKoon & Ratcliff, 1992; Moss et al., 1994; Plaut, 1995), but all these models differed from the one described here in fundamental ways—particularly in attention to connectivity and the use of itinerant dynamics with emergent metastable attractors. Thus, unlike most previous models, the model we describe can simulate whole trajectories of thought under the influence of priming, including the use of sequentially applied primes to shape this trajectory. We have already applied the present model to limited simulations of priming under specific conditions (Iyer et al., 2009, 2009), and further work is planned in this direction. As discussed above, the model may also be useful in investigating the effects of mental illness and dementia on semantic cognition. Finally, more complex versions of the model will also be used in future studies with large real-world datasets.

Acknowledgments

This work was supported in part by a National Science Foundation Human and Social Dynamics Program grant to Ali Minai (BCS-0728413), which includes support from the Deputy Director of National Intelligence, and by a National Science Foundation CreativeIT grant to Ali Minai (IIS-0855714). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors gratefully acknowledge helpful input from many colleagues—especially Vince Brown, Simona Doboli, Paul Paulus, Dan Levine, Wlodzislaw Duch, and Steve Bressler.

References

Abbott, L. F., & Regehr, W. G. (2004). Synaptic computation. *Nature*, *431*, 796–803.

Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*, 47–97.

Amit, D. (1989). *Modelling brain function*. Cambridge, UK: Cambridge University Press.

Ashby, F. G. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442.

Bales, M. E., & Johnson, S. B. (2006). Graph theoretic modeling of large-scale semantic networks. *Journal of Biomedical Informatics*, *39*, 451–464.

Barabasi, A. L. (2002). *Linked: the new science of networks*. Cambridge, MA: Perseus.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*, 509–511.

Bohland, J. W., & Minai, A. A. (2001). Efficient associative memory using small-world architecture. *Neurocomputing*, *38–40*, 489–496.

Bressler, S. L., & Tognoli, E. (2006). Operational principles of neurocognitive networks. *International Journal of Psychophysiology*, *60*, 139–148.

Brown, V., Tumeo, M., Larey, T., & Paulus, P. (1998). Modeling cognitive interactions during group brainstorming. *Small Group Research*, *29*, 495–526.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, *67*, 380–400.

Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends in Cognitive Sciences*, *7*, 354–361.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic priming. *Psychological Review*, *82*, 407–428.

Damasio, A. R. (1989). The brain binds entities and events by multiregional activations from convergence zones. *Neural Computation*, *1*, 123–132.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*, 499–505.

Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*, 179–229.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., et al. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biological Cybernetics*, *60*, 121–130.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*, 704–716.

Fauconnier, G., & Turner, M. (2003). *The way we think: conceptual blending and the mind's hidden complexities*. Basic Books.

Graziano, M., Polosecki, P., Shalom, D. E., & Sigman, M. (2011). Parsing a perceptual decision into a sequence of moments of thought. *Frontiers in Integrative Neuroscience*, *5*, 45. doi:10.3389/fnint.2011.00045.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*, 2554–2558.

Iyer, L. R., Doboli, S., Minai, A. A., Brown, V. R., Levine, D. S., & Paulus, P. B. (2009). Neural dynamics of idea generation and the effects of priming. *Neural Networks*, *22*, 674–686.

Iyer, L., & Minai, A. A. (2011). A neurodynamical model of context-dependent category learning. In *Proceedings of IJCNN 2011* (pp. 2975–2982).

Iyer, L. R., Minai, A. A., Doboli, S., Brown, V. R., & Paulus, P. B. (2009). Effects of relevant and irrelevant primes on idea generation: a computational model. In *Proceedings of IJCNN 2009* (pp. 1380–1387).

Iyer, L. R., Venkatesan, V., & Minai, A. A. (2010). Neurcognitive spotlights: configuring domains for ideation. In *Proceedings of WCCI 2010* (pp. 3026–3033).

Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colourful or noisy? attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 207–221.

Kohonen, T. (1989). *Self-organization and associative memory*. Berlin: Springer-Verlag.

Kruschke, J. K. (1992). Alcove: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22–44.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). Sustain: a network model of category learning. *Psychological Review*, *111*, 309–332.

Maki, W. S. (2003). Some properties of word pairs: word association, semantic distance, and lexical co-occurrence. Retrieved September 14, 2011. From psychonomic society web archive: http://www.psychonomic.org/archive/.

Maki, W. S. (2008). A database of associative strengths from the strength-sampling model: a theory-based supplement to the Nelson, Mcevoy, and Schreiber word association norms. *Behavior Research Methods*, *40*, 232–235.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*, 71–87.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.

Marupaka, N., & Minai, A. A. (2011). Connectivity and creativity in semantic neural networks. In *Proceedings of IJCNN 2011* (pp. 3127–3133).

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 310–322.

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1155–1172.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220–232.

Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science*, *6*, 413–427.

Motter, A. E., de Moura, A. P. S., Lai, Y. C., & Dasgupta, P. (2002). Topology of the conceptual network of language. *Physical Review E*, *65*, 065102(R).

Nelson, D. L., Bennett, D. J., Gee, N. R., Schreiber, T. A., & McKinney, V. (1993). Implicit memory: effects of network size and interconnectivity on cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 747–764.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The university of south florida word association, rhyme, and word fragment norms. http://www.usf.edu/freeassociation/.

Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, *105*, 299–324.

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*, 322–348.

Nelson, D. L., & Xu, J. (1995). Effects of implicit memory on explicit recall: set size and word frequency effects. *Psychological Research*, *57*, 203–214.

Nemeth, C. (1995). Dissent as driving cognition, attitudes, and judgments. *Social Cognition*, *13*, 273–291.

Newman, M. E. J. (2010). *Networks: an introduction*. Oxford University Press.

Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distribution and their applications. *Physical Review E*, *64*, 026118.

Newman, M. E. J., Watts, D. J., & Strogatz, S. H. (2002). Random graph models of social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 2566–2572.

Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: a cognitive model of idea generation in groups. *Personality and Social Psychology Review*, *3*, 186–213.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53–79.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? the representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*, 976–987.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In *Proc. of the 17th annual conference of the cognitive science society* (pp. 37–42). Hillsdale, NJ, Pittsburgh, PA: Lawrence Erlbaum.

Plotkin, R. (2009). *The genie in the machine: how computer-automated inventing is revolutionizing law and business*. Stanford University Press.

Poincaré, H. (1921). *The foundations of science*. Science Press.

Raaijmakers, J. G. W., & Shiffrin, R. (1981). Search of associative memory. *Psychological Review*, *88*.

Rabinovich, M., Volkovskii, A., Lecanda, P., Huerta, R., Abarbanel, H. D. I., & Laurent, G. (2001). Dynamical encoding by networks of competing neuron groups: winnerless competition. *Physical Review Letters*, *87*, 068102(4).

Ratcliff, R., & McKoon, G. (1994). Retrieving information from memory: spreading-activation theories versus compound-cue theories. *Psychological Review*, *101*, 177–184.

Schilling, M. A. (2005). A small-world network model of cognitive insight. *Creativity Research Journal*, *17*, 131–154.

Sigman, M., & Cecchi, G. A. (2002). Global organization of the wordnet lexicon. *Proceedings of the National Academy of Sciences*, *99*, 1742–1747.

Simonton, D. K. (1988). *Scientific genius: a psychology of science*. Cambridge University Press.

Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: the integration of product, person, and process perspectives. *Psychological Bulletin*, *129*, 475–494.

Simonton, D. K. (2010). Creative thought as blind-variation and selective-retention: combinatorial models of exceptional creativity. *Physics of Life Reviews*, *7*, 156–179.

Singer, W., Engel, A. K., Kreiter, A. K., Munk, M. H. J., Neuenschwander, S., & Roelfsema, P. R. (1997). Neuronal assemblies: necessity, signature and detectability. *Trends in Cognitive Sciences*, *1*, 252–260.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Reviews of Neuroscience*, *18*, 555–586.

Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, *2*, 145–162.

Starzyk, J. (2011). Mental saccades in the cognitive process. In *Proceedings of the 2011 international joint conference on neural networks*. (pp. 495–502).

Steyvers, M., & Tenenbaum, J. B. (2005). The large scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41–78.

Thaler, S. L. (1996a). Neural nets that create and discover. In *PC AI May/June* (pp. 16–21).

Thaler, S. L. (1996b). A proposed symbolism for network-implemented discovery processes. In *Proceedings of the 1996 world congress on neural networks*.

Tsuda, I. (2001). Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behavioral and Brain Sciences*, *24*, 793–847.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Science*, *5*, 244–252.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*, 229–239.

Wang, D. L., & Terman, D. (1995). Locally excitatory globally inhibitory oscillatory networks. *IEEE Transactions on Neural Networks*, *6*, 283–286.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*, 829–854.

Watts, D. J. (1999). *Small worlds: the dynamics of networks between order and randomness*. New Jersey: Princeton University Press.

Watts, D. J. (2003). *Six degrees: the science of a connected age*. New York: W.W. Norton.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, *393*, 440–442.

Yook, S. H., Jeong, H., & Barabasi, A. L. (2002). Modeling the Internet's large-scale topology. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 13382–13386.

Zucker, R. S., & Regehr, W. G. (2002). Short-term synaptic plasticity. *Annual Review of Physiology*, *64*, 355–405.

Please cite this article in press as: Marupaka, N., et al. Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. Neural Networks (2012), doi:10.1016/j.neunet.2012.02.004