IJCSS International Journal of Complex Systems in Science

Int. J. Complex Systems in Science
vol. **1** (2011), pp. 69–72

Navigation and Cognition in Semantic Networks

Javier Borge-Holthoefer[1,†] and Alex Arenas[1]

[1] *Department d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Catalonia, Spain*

Abstract. Semantic memory is the cognitive system devoted to storage and retrieval of conceptual knowledge. Empirical data indicate that semantic memory is organized in a network structure. Everyday experience shows that word search and retrieval processes emerge providing fluent and coherent speech, i.e. are efficient and robust. Nonetheless, links between pairs of words in semantic memory encode a rich variety of relationships, and not merely category membership. To extract this information, we schematize a process based on uncorrelated random walks from node to node, which converge to a feature vectors network. This mechanism forces the emergence of semantic similarity, which implicitly encloses category structure. Interestingly, the degradation of the original structure has a dramatic impact on the topology of semantic network, whereas the dynamics upon it evidence much higher resilience. We define this problem in the framework of percolation theory.

Keywords: information retrieval; complex networks; percolation; semantic impairment

1. Introduction

Semantic memory is the cognitive system where conceptual knowledge is stored. It can be suitably represented as a network, where nodes represent words and links between pairs of nodes stand for word–word relationships. A directed, weighted topology can be obtained from empirical Free Association (FA) data [1], which (see [2] for an interpretation in terms of complex networks). Taking these data as a proxy to the actual structure of semantic memory, several cognitive capacities can be studied, i.e. search, retrieval, category formation, lexical impairment, etc. Here we briefly tackle two related issues, namely category formation and disease-induced lexical impairment.

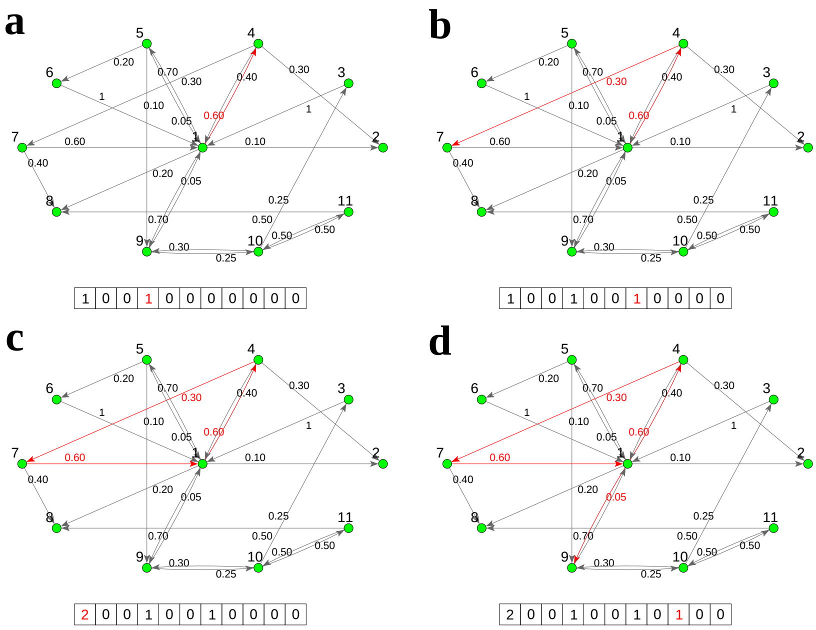

Figure 1: In RIM, the visits of a random walker starting at node i trigger the inheritance mechanism, which modifies the features vector of a node i. In the figure, a random walk of 4 steps (a–d) changes the vector of node 1 (color online).

2. The Random Inheritance Model

FA data reflects many possible ways by which two words can be related (semantic similarity, causal or functional relationship, etc.). The Random Inheritance Model (RIM) [2] explores whether it is possible to disentangle similarity relationships from general word association network (FA) by the navigation of the semantic network.

The idea is to simulate a naïve cognitive navigation on top of a general association semantic network to relate words with semantic similarity. Two words are considered semantically similar if they share features. The process can be schematized as uncorrelated *random walks* from node to node that propagate an *inheritance mechanism* among words, converging to a feature vectors network, see Figure 1 and Reference [2] for the complete explanation of the model.

The results obtained by RIM show macro-statistical coincidences (functional form of the distributions and descriptors) between real semantic simi-

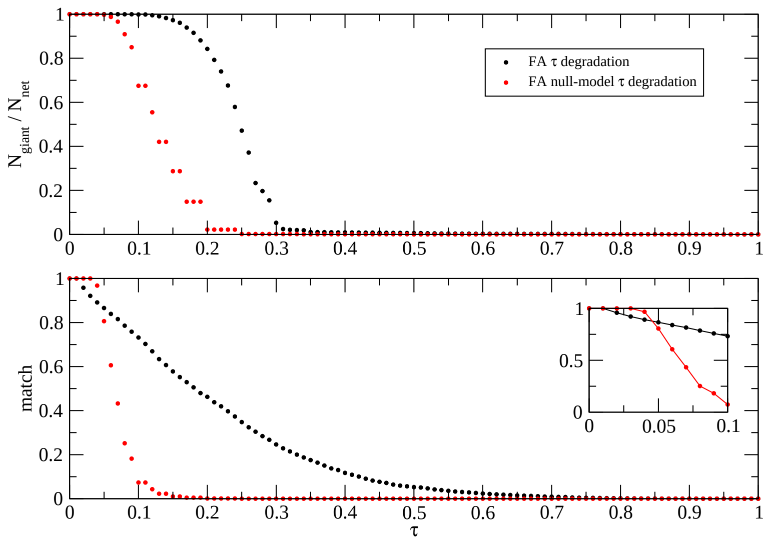

Figure 2: Upper: topological deterioration of FA as a function of τ. In black, results for the original FA structure. The same process of degradation has been applied to an unweighted version of FA understood here as a plausible null model, in red. Lower: RIM's resilience for the same structures.

larity data and the synthetic obtained network. Moreover, the model yields also significant success at the microscopic level, i.e. is able to reproduce to a large extent empirical relationships, see [2]. These results support the general hypothesis about implicit entangled information in FA, and also reveals a possible mechanism of navigation to recover feature information in semantic networks.

3. Semantic Memory robustness

In the previous section we have proposed a mechanism that drives the emergence of category structure. Now we turn to the characteristics of both the original topology and RIM dynamics under error. Literature on error and attack tolerance in complex networks [3, 4] typically model error as the failure (removal) of randomly chosen nodes/edges. Here, we redefine error in the context of cognitive systems. In this framework, it is more useful to consider error in terms of aging or disease, where the whole topology simultaneously decays in some way. By doing so, we capture the degrading behavior of aging and/or disease, which differs from attack (there is no selective action) and from error (which affects only one node/edge at a time). For the sake of clarity, we refer to error in the cognitive framework as degradation.

Degradation assumes that links are increasingly damaged. At a given threshold τ, every link (i, j) in FA with a $\omega_{ij} \leq \tau$ is removed. This process is performed with values $0 \leq \tau \leq 1$. Also, for each value τ, we monitor both topological and dynamical properties of the resulting network. On one hand, the size of the giant component of the degraded structure is measured. On the other, RIM is used to find a similarity matrix on the degraded structure, and the result is compared to the non-degraded RIM, i.e. RIM's result at $\tau = 0$ ("Match" axis in Figure 2). Figure 2 shows the results for both topological deterioration and dynamical resilience. The behavior of RIM's dynamics appears to be very sensitive to degradation even at very low values of τ. This implies that lexical impairment can appear at early stages of semantic memory aging or disease degradation. Interestingly, however, RIM's degradation is much slower than the topological one. At $\tau \approx 0.3$, FA structure is disintegrated already, whereas RIM can still recover as much as a 20% of its original content. RIM's results do not vanish up to $\tau \approx 0.6$.

4. Conclusions

We have introduced RIM as a plausible cognitive dynamics to extract the category structure backbone from a general semantic relations context (FA). Once this information is available we study to what extent a progressive degradation of the topological structure affects lexical performance. To this end, we follow the line of percolation theory in complex network with some modifications. Results indicate that linguistic performance is severely affected by semantic memory degradation, on the other hand such performance is still significantly effective way beyond topological disintegration.

References

[1] D.L. Nelson, C.L. McEvoy and T.A. Schreiber, *http://www.usf.edu/FreeAssociation*, (1998).

[2] J. Borge-Holthoefer and A. Arenas, *European Physical Journal B*, **74**, 265–270 (2010).

[3] R. Albert, H. Jeong and A.L. Barabási *Nature*, **406**, 378–382 (2000).

[4] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin *Physical Review Letters*, **85(21)** (2000).