UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
U-INVITE: Estimating Individual Semantic Networks from Fluency Data

Permalink
https://escholarship.org/uc/item/1st4x5c3

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Zemla, Jeffrey C.
Kenett, Yoed N.
Jun, Kwang-Sung
et al.

Publication Date
2016

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

U-INVITE: Estimating Individual Semantic Networks from Fluency Data

Jeffrey C. Zemla (jzemla@brown.edu)
Yoed N. Kenett (yoed_kenett@brown.edu)
Department of Cognitive, Linguistic, and Psychological Sciences
Brown University, Providence, RI 02912 USA

Kwang-Sung Jun (kjun@discovery.wisc.edu)
Wisconsin Institute for Discovery
University of Wisconsin-Madison, Madison, WI 53706 USA

Joseph L. Austerweil (joseph_austerweil@brown.edu)
Department of Cognitive, Linguistic, and Psychological Sciences
Brown University, Providence, RI 02912 USA

Abstract

Semantic networks have been used extensively in psychology to describe how humans organize facts and knowledge in memory. Numerous methods have been proposed to construct semantic networks using data from memory retrieval tasks, such as the semantic fluency task (listing items in a category). However these methods typically generate group-level networks, and sometimes require a very large amount of participant data. We present a novel computational method for estimating an individual's semantic network using semantic fluency data that requires very little data. We establish its efficacy by examining the semantic relatedness of associations estimated by the model.

Keywords: semantic networks; memory retrieval; fluency; random walk; probabilistic modeling

Introduction

Semantic memory is the system of memory that stores concepts and facts. Although the way in which semantic memory is organized into categories and subcategories remains an open question (Jones, Willits, & Dennis, 2015), one common approach is to represent it as a network comprised of nodes (a word or concept) and edges between nodes that signify that the two concepts are associated. However, how do we estimate a given individual's semantic network?

A growing body of work has related statistics of semantic networks (e.g., centrality) to cognitive phenomena, such as language development, memory retrieval and creative thinking (Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013; Hills, Todd, Lazer, Redish, & Couzin, 2015). Most of this work has analyzed aggregated group-based networks, which cannot be used to understand individual differences. Currently, only one study has examined individual differences in semantic networks (Morais, Olsson, & Schooler, 2013). Here, we present a novel probabilistic method to estimate an individual's semantic memory structure efficiently using data from a semantic fluency task.

The semantic fluency task (listing of items in a category) has a long history in cognitive psychology (Henley, 1969). Typical subjects show a distinct behavioral pattern in this task, reporting items in clusters (sub-categories) and switching to new clusters when subsequent items are hard to retrieve (Troyer, Moscovitch, & Winocur, 1997). For instance, when typical subjects list animals, they may list several farm animals before switching to zoo animals. This clustering and switching behavior has been used to make inferences about the cognitive processes and representations underlying search through semantic memory.

Abbott, Austerweil and Griffiths (2012, 2015) proposed a model of semantic memory retrieval that accounts for this clustering and switching behavior (though see Hills, Jones, & Todd, 2012 for an alternative model that also accounts for this behavior). Given a semantic network, data are generated by taking a censored random walk on that network: Starting from a category's node, their model moves over random edges, emitting the labels of any nodes (e.g., "turkey") in its path if they are in the target category and have not been visited previously. The result is a fluency list that contains no duplicate items and is arranged by the order in which the items were first encountered. Jun et al. (2015) proposed a computational method based on this process to infer a semantic network from fluency data. Their method, initial-visit emitting random walk (INVITE) is based on the principle that multiple fluency lists from the same network can be used to infer that semantic network.

In this paper, we build on the INVITE model to develop a novel method for estimating an individual's semantic network.. We begin by presenting a few possible methods to estimate semantic networks, including INVITE and our approach. Next we evaluate these approaches and show that our approach can efficiently recover a network from simulated fluency data. Finally, we present an experiment where we collected fluency data from participants and examined the semantic similarity of edges in networks generated by the different approaches.

Estimating Semantic Networks

To estimate a network from fluency lists, we assume items are retrieved according to a censored random walk on a network (Abbott et al., 2015) and invert this process using Bayes' rule. Formally, let **G** be the participant's semantic network, $\vec{X}^m$ be the *m*th list produced from the participant (a censored random walk from the network), $\vec{S}^m$ be the inter-item response times (IRTs) of the *m*th list (so S_k^m is the time between response k and k-1 in list *m*), $\vec{Z}^m$ be the *m*th uncensored random walk on **G** (all category members visited by the random walk regardless of whether they were previously said) and $c(\vec{Z}^m)$ be a censoring function applied to the uncensored random walk such that it returns the censored walk, (i.e., $c(\vec{Z}^m) = \vec{X}^m$). We assume that each IRT S_k^m is gamma-distributed (e.g., Luce, 1986) with parameters $\tau_k^m - \tau_{k-1}^m$ and β. The former parameter reflects the number of censored items between two unique items in the *m*th uncensored list, where τ_k^m is the index of the *k*th unique item reported in the *m*th uncensored list. β is a parameter that controls the amount of variability in response times. Intuitively, the first parameter increases the expected IRT (so as the number of censored items between two uncensored items increases, the expected IRT increases) and β controls the variance (see Figure 1).

We examine this model as well as a restricted model that does not include response times, and a naïve random walk model that assumes no censoring occurs.

The Naïve Random Walk Model

The naïve random walk procedure (RW) ignores the censoring procedure and places edges between all successive items in every fluency list as if there were no censored items. For example, if we have a single list "*dog*, *cat*, *mouse*", our network would consist of three nodes and two edges, *dog-cat* and *cat-mouse*. When few lists are available, the RW procedure is a close fit to the most likely network. However, when many lists are available, a network estimated using this procedure quickly becomes over-connected, resulting in a network that contains many false edges. The RW procedure is inconsistent, meaning that it is not guaranteed to converge and, in fact, it will typically become less accurate as the number of lists increases.

The INVITE Model

Jun et al. (2015) proposed a method to invert the generative process used by Abbott et al. (2015): Given a participant's semantic fluency data were produced by a random walk on a network, what is the most probable network? Given M fluency lists, each denoted as $\vec{X}^m = (X_1^m \dots X_{Nm}^m)$, we seek a network **G** that maximizes the likelihood of the data:

$$\prod_{m=1}^{M} \prod_{k=1}^{N_m - 1} \mathbb{P}(X_{k+1}^m | X_{1:k}^m) \qquad (1)$$

where N_m denotes the length of the *m*th censored list and X_k^m denotes the *k*th item from the *m*th list. Hereafter, we remove

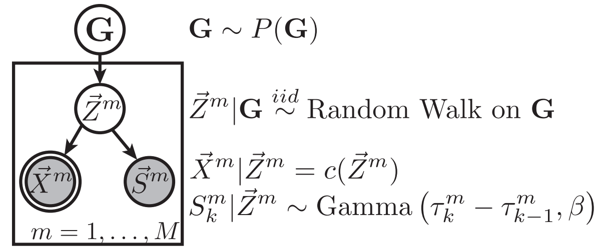

Figure 1: Left: Graphical model representing our computational method. The box is a plate, which means that the variables are copied and conditionally independent given the network **G** for each list *m* from one to the total number of lists *M*. The shaded nodes are observed. The double circle denotes a deterministic function. Right: The generative process for the model. iid means independent and identically distributed.

this superscript for readability when it is clear from context. The key to INVITE is to generate each item in a fluency list from a different random walk – one that treats visited nodes as transient and unvisited nodes as absorbing. To form each random walk, we re-arrange the states in the transition matrix of **G** so that the rows and columns are in list order, e.g., G'_{12} denotes a transition from X_1 to X_2:

$$\mathbf{G}' = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

where **Q** denotes transitions between previously emitted items (transient states), **R** denotes transitions from previously emitted items to novel items (absorbing states), and **0** and **I** (the identity matrix) ensure the random walk is absorbing. **G'** is updated after each step in a list and reconfigured after each list.

Thus, we calculate $\mathbb{P}(X_{k+1}|X_{1:k})$ as the probability of starting at X_k and being absorbed by X_{k+1} given transient states $X_{1:k}$ and absorbing states $X_{k+1:Nm}$. This is computed using the fundamental matrix (Doyle & Snell, 1984) of **G'**, $\mathbf{N}=(\mathbf{I}-\mathbf{Q})^{-1}$, where N_{ij} denotes the expected number of times a walk starting from state i visits state j before being absorbed. Thus,

$$\mathbb{P}(X_{k+1}|X_{1:k}) = \begin{cases} \sum_{i=1}^{k} N_{ki} R_{i1} & \text{if } \mathbf{N} \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

The U-INVITE Model

Our model, U-INVITE, improves performance of the INVITE model given a small number of lists. It does so by extending INVITE in two ways: (1) by using the time between responses (IRTs), and (2) assuming that the network is undirected and unweighted (the probability of transiting to any connected node is equal).

Inter-item Response Times As shown in Jun et al. (2015), INVITE works particularly well when the number of lists is large. When the number of lists is small, there may not be enough information in the order of the items to accurately estimate the network. One way we resolve this is by incorporating IRTs into the emission probability. Rather than maximizing Equation 1, we maximize the following equation:

$$\prod_{m=1}^{M}\prod_{k=1}^{N_m-1}\sum_{r=1}^{\infty}\mathbb{P}(X_{k+1}^{m}\text{ in r steps}|X_{1:k}^{m})^{\theta}\mathbb{P}(S_{k+1}^{m}|r,\beta)^{(1-\theta)}$$

That is, we weight the probability of being absorbed by X_{k+1} in r steps by the probability of observing an IRT of S_{k+1} given an r-step walk. We calculate the probability of observing an IRT given r steps using a gamma distribution parameterized by β. We add an additional free parameter θ to adjust the influence of IRTs on the model, i.e., when θ is 1 the model ignores IRT information. Although this equation contains an infinite summation, we have found that limiting this to $r = 20$ works as an efficient approximation in practice, as most chains are absorbed by the next state in fewer than 20 steps. Rather than computing the probability of being absorbed by X_{k+1} in the limit, we compute:

$$\mathbb{P}(X_{k+1}\text{ in r steps}|X_{1:k})=\sum_{i=1}^{k}\left(\mathbf{Q}^{r-1}\right)_{ki}R_{i1}$$

We assume a uniform prior for $\mathbf{G}$, and that $\mathbb{P}(X_1^m|\mathbf{G})$ is uniformly distributed for all M lists.

Unweighted Networks and the Search Procedure In addition to timing information, we include additional constraints to estimate networks efficiently: We assume that the random walk is unweighted and undirected. Although these assumptions may seem psychologically unrealistic, Abbott et al. (2015) found that both weighted and unweighted semantic networks captured human performance in semantic fluency tasks well. Whether human semantic networks are unweighted or weighted is an unsettled question and orthogonal to the purpose of our paper (a method for estimating weighted networks with IRT information could be created by deriving a MLE estimator without constraints on the transition matrix, as in Jun et al., 2015). Its strength enables us to estimate networks efficiently from censored lists. The original INVITE allows weighted edges, adding additional degrees of freedom that need to be inferred. Further, the transition matrix inferred by INVITE is fully-connected; to convert it into a network that is not fully-connected would require an additional thresholding process (where estimated edge weights lower than an additional threshold parameter are removed from the final network and then appropriately normalized). For these reasons, it is difficult to compare networks constructed by INVITE and U-INVITE, and we do not provide a direct comparison of the algorithms in this paper.

To find the network that maximizes the likelihood of the data, we use a stochastic search procedure with smart initialization. Using an initial network constructed with the RW procedure, we toggle one or more edges and compute the new network's probability, accepting the change when the new network is more probable given the data. We favor toggling edges that connect two items present in multiple lists, as these "hub nodes" have a larger effect on the network's posterior probability. Specifically, we set a fixed probability (P_{hub}=.8) that we toggle an edge that connects two hub nodes, or otherwise toggle an edge at random. We also favor toggling one edge at a time, as the probability of producing a network that has zero probability (cannot produce the data) increases rapidly as the number of simultaneous edge changes is increased. At each update, we toggle 1+D edges simultaneously where D is sampled from a Geometric distribution with P_{geom}=.2. These free parameters affect only the time to convergence, and ensure that the search procedure will converge in the limit. We run this procedure repeatedly until we have tried 1500 updates without finding a network with a higher likelihood. We found this stopping criterion to be robust for the toy networks estimated in this article.

Simulations

Varying the Number of Lists

We used simulated data to estimate the accuracy of four different models as a function of the number of fluency lists used to fit the network. We compared RW, U-INVITE, and U-INVITE with IRTs. We report results using two possible values of θ in the IRT method: 0.5 (the IRT5 model) and 0.9 (the IRT9 model).

We generated 10 toy small-world networks, consisting of 15 nodes each, using the Watts-Strogatz procedure (Watts & Strogatz, 1998). Previous literature has suggested that human semantic networks are small-world like (e.g., Borge-Holthoefer & Arenas, 2010), being highly clustered yet having a low shortest path length between any two nodes. We chose parameters for the Watts-Strogatz procedure to generate networks that were roughly comparable in node degree and clustering coefficient to what has been reported previously for human semantic networks (Steyvers & Tenenbaum, 2005). Our toy networks had an average node degree of 4 and a mean clustering coefficient of 0.29.

We varied the number of fluency lists used to estimate the network from 2 to 35. Lists were generated by starting at a random node in the network and taking a random walk until all of the nodes were traversed, then extracting only the first visit of each node from the list. Each list was truncated to roughly 70% of its length, or 11 items, with the restriction that each node in the network is traversed at least once in the set of lists. This truncation process mimics human-generated data reported later (i.e., each list contains approximately 70% of the total items listed by a participant). Simulated IRTs were generated from a gamma distribution, using the number of steps between two items in a walk and β=1.1 as parameters.

We calculated the cost of each reconstructed network using Hamming distance, or the number of edges that would need to be added or removed to convert it to the original network. The results, shown in Figure 2, demonstrate that U-INVITE does converge to the original network, though incorporating IRTs can lead to convergence with fewer lists. While the IRT5 model performs reasonably well, we found that it was outperformed by the IRT9 model, which assigns a higher weight to the item order than to the IRTs.

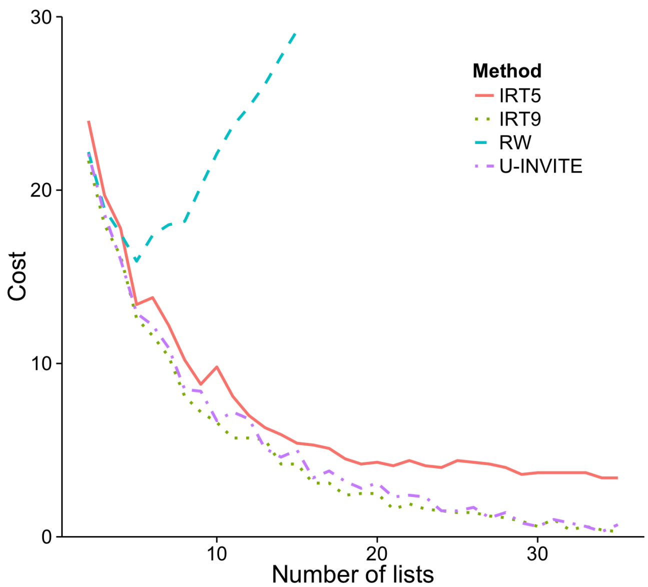

Figure 2: As the number of fluency lists increases, our methods tend toward zero error. Not shown: the RW method increases linearly to about 50 by 35 lists.

Model Comparison Given Three Lists

We conducted an additional simulation using only three fluency lists to examine whether IRTs improve network estimation when only a small number of lists are used. Using the same procedure as above, we generated 300 toy networks with 15 nodes each, and reconstructed each network using each of the four methods. A one-way repeated measures ANOVA was conducted to examine the effect of the different methods on the cost of estimating the original network (Table 1). This analysis revealed a significant main effect of method, $F(3, 897) = 26.13$, $p < 0.001$, $\eta^2 = .08$. Post-hoc analyses revealed that this main effect was due to the IRT9 model outperforming the other three models (all p's < 0.001). No significant differences were found between the average cost of the RW, U-INVITE, and IRT5 models.

We classified the edges in the reconstructed networks to indicate whether an edge was present in both the original and reconstructed network (Hit), in the original but not the reconstructed network (Miss), or in the reconstructed but not the original network (False Alarm). A one-way repeated measures ANOVA was conducted for Hits/Misses, $F(3, 897) = 262.06$, $p < 0.001$, $\eta^2 = .47$, as well as False Alarms, $F(3, 897) = 415.79$, $p < 0.001$, $\eta^2 = .51$. We found that compared to U-INVITE, the IRT9 model contains significantly more hits (and fewer misses). However, we found no difference in the number of false alarms. This indicates that incorporating IRTs improves upon U-INVITE by accurately detecting more genuine edges, while keeping the number of false alarms constant. In contrast, the RW method generates substantially more false alarms compared to U-INVITE and IRT9. Finally, the IRT5 model resulted in the highest amount of hits and fewest misses, but also the most false alarms. In future work, we will explore why this is the case and how to weight IRT and order information optimally. Next, we compare the different methods on real behavioral results from a semantic fluency task where participants generate multiple fluency lists.

Table 1: Results of estimating networks from three lists. 300 networks were generated and each method was used to estimate these toy networks. Values denote average scores (standard deviation in parentheses).

Measure	RW	U-INVITE	IRT5	IRT9
Cost	19.2 (3.3)	18.9 (3.6)	19.1 (3.7)	18.2 (3.5)
Hits	17.7 (1.8)	16.3 (2.0)	18.3 (1.9)	16.9 (2.0)
Misses	12.3 (1.8)	14 (2.0)	11.7 (1.9)	13.0 (2.0)
False Alarms	6.9 (1.9)	5.0 (2.2)	7.5 (2.4)	5.0 (2.1)

Experiment: A repeated semantic fluency task

Methods

Participants We recruited twenty participants from Amazon Mechanical Turk (11 male, mean age 31.75) who were located in the United States.

Procedure Participants were given a category label (e.g., *animals*) and asked to generate as many items from that category as possible in three minutes. Each participant completed nine lists in total, three for each of three categories (animals, fruits, and vegetables). The order of the lists was pseudo-randomized so that each triad of lists contained one of each category, and participants never completed the same category twice in succession.

Each response was hidden from view after it was entered to reduce cueing effects from previously entered items. Participants were instructed to list each item no more than once within a list, but that they could repeat themselves on subsequent lists.

Results

We present the results solely from the *animal* category, which generated the most responses. Twenty participants generated 280 unique animals in total (average 54.5 per participant and 33.7 per list). Participants were largely successful at avoiding repetitions, repeating fewer than one animal per list on average. All repetitions were removed from the data set prior to analysis.

To validate our method, we examined the similarity between connected nodes using the BEAGLE lexical semantic database (Jones & Mewhort, 2007). The database estimates the semantic similarity between two words from their statistical co-occurrence in a large corpus of text. For example, *dog-cat* has a high BEAGLE similarity whereas *dog-toad* has a low BEAGLE similarity.

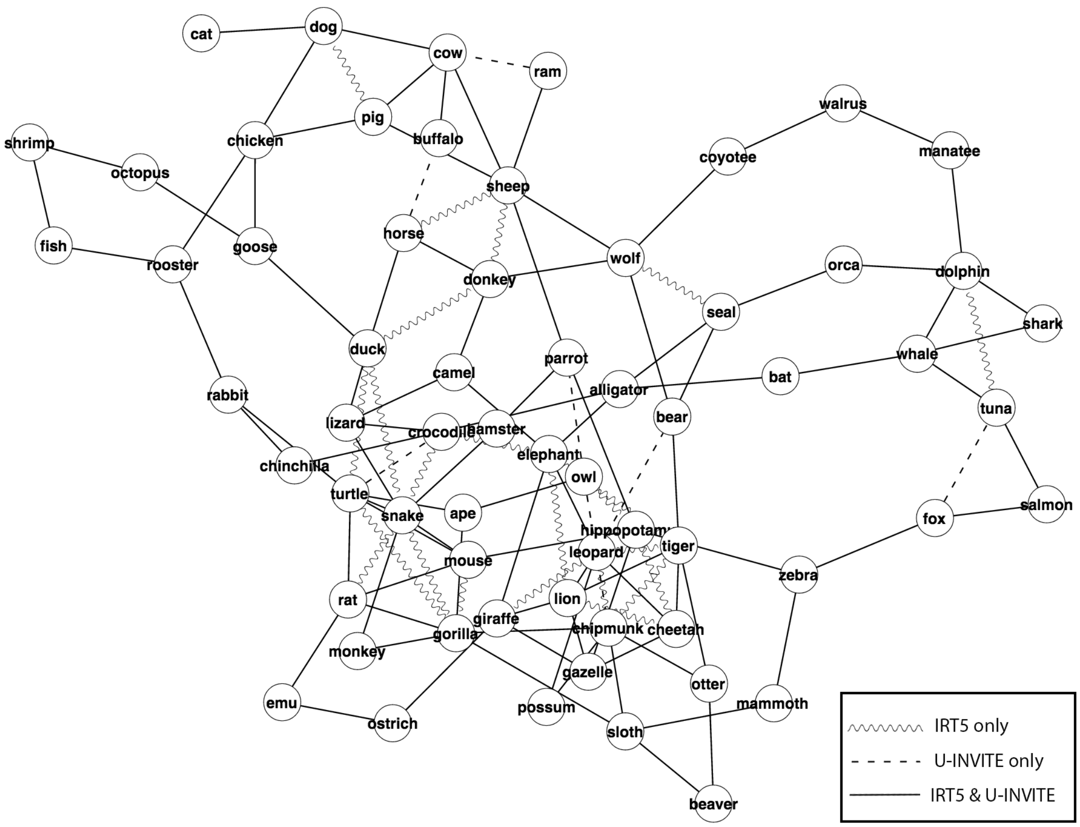

Figure 3: Network reconstruction with U-INVITE and IRT5 models of the semantic network of a single participant. Edge style denotes model success at estimating edges: Solid line: edges estimated by both methods; Dashed line: edges estimated only by the U-INVITE model; Sinusoidal line: edges estimated only by the IRT5 model.

We computed the BEAGLE value for all edges in each network, except those that could not be computed because one of the nodes was not in the BEAGLE database (accounting for 6.4% of edges). For each participant, we also calculated the BEAGLE value that would be expected between a random pair of nodes in the network (i.e., the average BEAGLE value across all possible edges in the network). We then subtract this score from the BEAGLE value of each edge in the network to generate a BEAGLE difference score (BDS), where positive scores indicate the two connected animals are more similar than would be expected by chance.

All five methods generated sensible networks[1]: the average BDS of all edges for each participant and each method (80 networks) was higher than would be expected by chance. U-INVITE generated networks that were similar to the RW method. Across all twenty participants, U-INVITE added zero edges compared to the RW method, and removed only 39 edges (roughly 2.4% of all edges). This is probably because U-INVITE needs more data to make accurate estimates. Jun et. al (2015) found that for toy networks, INVITE was typically a poor estimator of the true network when the number of lists was small.

Compared to U-INVITE, the IRT9 model added 23 edges (1.4%) and removed 22 edges (1.4%). We calculated the average BDS (per participant) of edges present in IRT9 but not in U-INVITE, and found that these edges had a BDS significantly greater than expected by chance, indicating that the added edges connect nodes that are semantically similar, $M_{BDS} = .068$, $t(12) = 2.21$, $p = .047$. However, we also found that edges present in U-INVITE but not in IRT9 were more similar than expected by chance, $M_{BDS} = .063$, $t(10) = 3.23$, $p = .009$. There was no statistical difference between the average BDS of edges added compared to edges removed ($p = .89$).

The IRT5 model made substantial changes to the network compared to U-INVITE (see Figure 3 for an example). Across all participants, the IRT5 method added 486 edges (30.1%) that were not present in the U-INVITE network and removed 150 edges (9.3%) of the edges that were present in the U-INVITE network. The edges added by IRT5 have an average BDS significantly greater than expected by chance, $M_{BDS} = .015$, $t(19) = 2.91$, $p = .017$. However, as with IRT9, we also found the reverse to be true: Edges that were removed from the U-INVITE model have a higher BDS score than was expected by chance, $M_{BDS} = .025$, $t(18) =$

[1] Networks for each participant and each method are available online at http://research.clps.brown.edu/austerweil/UINVITE16/

4.34, $p < .001$. Comparison between the edges removed and edges added showed no statistical difference ($p = .22$).

Conclusions

Advancements in network science and probabilistic modeling enable scientists to investigate how the structure of semantic memory contributes to language development, creativity and intelligence, and memory retrieval (De Deyne, Kenett, Anaki, Faust, & Navarro, in press). However, current research has been limited to group analyses, which cannot account for individual differences. Our method estimates an individual's semantic memory structure based on multiple semantic fluency responses.

Our approach extends that of Jun et al. (2015) by constraining the estimated networks to be unweighted and undirected, and incorporating response time information. We found that our method accurately recreated small-world networks, which have consistently been found to resemble human networks (Borge-Holthoefer & Arenas, 2010). Further modifications to our model may improve its accuracy. For instance, we may use a more realistic response time function other than a gamma distribution, such as the ex-Gaussian distribution (Heathcote, Popiel, & Mewhort, 1991). We also plan to examine more realistic process models (e.g., an imperfect censoring function would allow us to model perseverations in semantic retrieval). Finally, we plan to examine how to weight IRT information optimally and perform additional validations of our method.

Developing methods to estimate an individual's network representation from fluency data has great potential across cognitive science. They will allow us to examine individual differences in semantic networks and relate them to neurocognitive variables that affect memory search and executive functions in typical and clinical populations (Faust & Kenett, 2014). For instance, Alzheimer's and semantic dementia patients show marked disruption in performance on a semantic fluency task (Rohrer, Salmon, Wixted, & Paulsen, 1999). We hope that our method can be used to improve our understanding of impaired and unimpaired cognitive search.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2012). Human memory search as a random walk in a semantic network. *Advances in Neural Information Processing Systems, 25*, 3050-3058.

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Random walks on semantic networks can resemble optimal foraging. *Psychological Review, 122*(3), 558-569. doi:10.1037/a0038693

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in Cognitive Science. *Trends in Cognitive Sciences, 17*(7), 348-360. doi:10.1016/j.tics.2013.04.010

Borge-Holthoefer, J., & Arenas, A. (2010). Semantic networks: Structure and dynamics. *Entropy, 12*(5), 1264-1302.

De Deyne, S., Kenett, Y. N., Anaki, D., Faust, M., & Navarro, D. J. (in press). Large-scale network representations of semantics in the mental lexicon. In M. N. Jones (Ed.), *Big data in cognitive science: From methods to insights*: Psychology Press.

Doyle, P. G., & Snell, J. L. (1984). *Random walks and electric networks*.

Faust, M., & Kenett, Y. N. (2014). Rigidity, chaos and integration: Hemispheric interaction and individual differences in metaphor comprehension. *Frontiers in Human Neuroscience, 8*(511), 1-10.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: an example using the Stroop task. *Psychological Bulletin*, *109*(2), 340-347.

Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 176-184.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119*(2), 431-440. doi:10.1037/a0027373

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences, 19*(1), 46-54.

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114*(1), 1-37. doi:10.1037/0033-295X.114.1.1

Jones, M. N., Willits, J., & Dennis, S. (2015). Models of semantic memory. In J. Busemeyer & J. Townsend (Eds.), *Oxford Handbook of Mathematical and Computational Psychology* (pp. 232-254).

Jun, K.-S., Zhu, X., Rogers, T. T., Yang, Z., & Yuan, M. (2015). Human memory search as initial-visit emitting random walk. *Advances in Neural Information Processing Systems (NIPS)*.

Luce, R. D. (1986). *Response times* (No. 8). Oxford University Press.

Morais, A. S., Olsson, H., & Schooler, L. J. (2013). Mapping the structure of semantic memory. *Cognitive Science, 37*(1), 125-145. doi:10.1111/cogs.12013

Rohrer, D., Salmon, D. P., Wixted, J. T., & Paulsen, J. S. (1999). The disparate effects of Alzheimer's disease and Huntington's disease on semantic memory. *Neuropsychology, 13*(3), 381-388.

Steyvers, M., & Tenenbaum, J. B. (2005). The large scale structure of semantic networks: Statistical analysis and a model of semantic growth. *Cognitive Science, 29*(1), 41-78.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138-146.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(4), 440-442.