New Ideas in Psychology xxx (2017) 1–7

Contents lists available at ScienceDirect

New Ideas in Psychology

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

In defense of spatial models of semantic representation☆

Michael N. Jones [a,*], Thomas M. Gruenenfelder [a], Gabriel Recchia [b]

[a] *Indiana University, USA*
[b] *University of Cambridge, United Kingdom*

ARTICLE INFO

Article history:
Received 6 January 2017
Received in revised form
22 August 2017
Accepted 25 August 2017
Available online xxx

ABSTRACT

Recent semantic space models learn vector representations for word meanings by observing statistical redundancies across a text corpus. A word's meaning is represented as a point in a high-dimensional semantic space, and semantic similarity between words is quantified by a function of their spatial proximity (typically the cosine of the angle between their corresponding vector representations). Recently, Griffiths, Steyvers, and Tenenbaum (2007) demonstrated that spatial models are unable to simulate human free association data due to the constraints placed upon them by metric axioms which appear to be violated in association norms. However, it is important to note that free association data is the product of a retrieval process operating on a semantic representation, and the failures of spatial models are likely be due to mistaking the similarity metric (cosine) for an appropriate process model of the association task—cosine is not what people do with a memory representation. Here, we test the ability of spatial semantic models to simulate association data when they are fused with a simple Luce choice rule to simulate the process of selecting a response in free association. The results provide an existence proof that spatial models can produce the patterns of data in free association previously thought to be problematic for this class of models.

1. Introduction

A longstanding belief in theories of lexical semantics, dating back at least to Osgood (1952) is that words can be represented as points in a multidimensional semantic space. Similarity between word meanings is then defined as some function of their distance in space. This classic notion of mental space has had an obvious impact on modern computational semantic space models, such as Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). Models such as LSA borrow techniques from linear algebra to infer the semantic representation for words from their contextual co-occurrences in linguistic corpora. In the resulting space, a word's meaning is represented by a vector over latent dimensions. Inter-word similarity is based on Euclidean geometry: Words that are more similar are more proximal in the learned space. Virtually all distributional models of semantic memory adhere to the spatial notion of semantics (for a review, see Jones, Willits, & Dennis, 2015), including recent popular neural embedding models (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013).

In contrast to spatial models, the popularity of probabilistic models of cognition has led to the development of Bayesian models of semantic representation, such as the LDA-based Topic models explored by Griffiths et al. (2007). In a Topic model, a word's representation is a probability distribution over latent semantic "topics." When a word is processed, its semantic representation is the predicted probability across latent topics. Hence while LSA represents a word as a point in high-dimensional space and requires a spatial metric of similarity between two words, a Topic model represents a word as a probability distribution and computes the association between words as the probability of one word given the other. This allows Topic models to make very different predictions depending on which word is being conditioned upon, in contrast to LSA in which similarity is identical regardless of which word is "first." In addition, the issue of whether humans represent meaning as a coordinate in space or as a conditional probability is a fundamental question in cognitive science, and has implications for downstream models that make use of these representations.

Tversky (1977) has noted that spatial models must respect several metric axioms. Firstly, in a metric space the distance between a point and itself must be zero by any Euclidean metric,

☆ This work was supported by NSF BCS-1056744 and IES R305A150546.
* Corresponding author. Department of Psychological and Brain Sciences, Indiana University, Bloomington, USA.
E-mail address: jonesmn@indiana.edu (M.N. Jones).

http://dx.doi.org/10.1016/j.newideapsych.2017.08.001

$d(x,x) = 0$ (*non-negativity*). Secondly, distance must respect *symmetry*: $d(x,y) = d(y,x)$. Thirdly, distance must respect the *triangle inequality*: If x and y are proximal and y and z are proximal, then x and z are likely to be proximal points as well (specifically, $d(x,z) \leq d(x,y) + d(y,z)$). As Tversky and Gati (1982) have demonstrated, human judgments of similarity routinely violate these axioms—specifically, symmetry and the triangle inequality. Tversky used human violations of the metric axioms to argue against spatial models of similarity, and instead proposed an additive feature comparison model. The spatial debate, however, has a long history in cognitive science, with Tversky's work being followed by explanations of how metric spaces could produce violations of metric axioms (e.g., Krumhansl's (1978) notion of density or Holman's (1979) similarity and bias model).

Griffiths et al. (2007) note that human free association norms also violate metric axioms, making them problematic for semantic space models such as LSA. In a free association task, the participant is asked to respond to a cue word with the first associated word that comes to mind (Nelson, McEvoy, & Schreiber, 2004). Word association norms contain a significant number of asymmetric associations: For example, the probability of generating *baby* as a response to *stork* as a cue is much greater than the reverse. Part of this effect is due to a bias to respond with a high frequency target independent of the cue, but part appears to be due to some sort of asymmetry in the computation of similarity. In addition, word association norms contain apparent violations of the triangle inequality axiom: To use the example from Griffiths et al., *asteroid* is strongly associated with *belt*, and *belt* is strongly associated with *buckle*, but *asteroid* and *buckle* have no association. Finally, Griffiths et al (see also Steyvers & Tenenbaum, 2005). have demonstrated that association norms contain neighborhood structure that is incompatible with spatial models. If one constructs an associative network with nodes representing words and connecting edges based on nonzero association probabilities, the resulting networks are *scale-free*: they have power law degree distributions and high clustering coefficients.[1]

Griffiths et al. (2007) note, however, that probabilistic representations are not subject to the same metric restrictions as spatial representations, and they provide an elegant demonstration of how Topic models can naturally account for the qualitative nature of violations in asymmetry and the triangle inequality that LSA cannot. Griffiths et al. further demonstrate that while LSA (based on a thresholded cosine) cannot reproduce the scale-free and small-world network structure seen in word association norms, this structure naturally emerges in a Topic model.

However, it is important to note that an observable behavior such as free association is the product of a cognitive process operating on a memorial representation (Anderson, 1978; Estes, 1975). This notion is ubiquitous in cognitive science. For example, Nosofsky (1986) uses a spatial representation of stimuli, but the complex classification behavior of his model is the result of applying a simple choice rule to this spatial representation, not spatial distance itself. Similarly, semantic space models are models of memory structure; the structural model should not be expected to simulate a complex behavior like memory retrieval without the benefit of a process account to explain how the memory structure is used in a particular task. While the cosine between two word vectors is often used as a measure of their semantic similarity, it is a measure of the similarity of memory structures rather than an appropriate process model of the task—*a* cosine *is not what people do in a task*, and should not be used as an estimate of behavioral data (see Jones, Hills, & Todd, 2015). A cosine, or similar metric, should be the input to a process model if one is interested in simulating behavioral data. This also enhances the models' generalizability across different tasks that tap semantic structure, and is particularly appealing given the low correlation in responses between different tasks thought to utilize the same semantic structure (Maki & Buchanan, 2008), and the fact that different semantic space models give the best fit to different behavioral tasks even though all tasks are thought to tap the same semantic memory structure (Mandera, Keuleers, & Brysbaert, 2017).

Griffiths et al. (2007, p. 224) imply that a "more complex" spatial metric based on LSA (similar to Nosofsky's 1986, 1991 use of a similarity-choice function) could potentially account for the metric axiom violations in association norms. We return to the issue of complexity with regard to spatial and probabilistic models in the discussion. The bulk of this paper will be focused on evaluating this suggestion by fusing spatial semantic models with a parameter-free version of Luce's (1959) similarity-choice model to evaluate their ability to account for the problematic data identified by Griffiths et al. In doing so, we provide an existence proof that semantic space models can indeed produce asymmetries, violations of the triangle inequality, and scale-free network structure with an appropriate process rule. It is premature to reject spatial models of semantic representation based on violations of metric axioms in association data.

2. A generic spatial choice model

In this paper, we evaluate the application of Luce's (1959) classic choice rule to simulate the cognitive process involved in the task of free association when applied to three (metric) semantic space models, gradually increasing in complexity. Although similarity and distance in the semantic spaces respect the metric axioms, the behavior of the choice rule applied to these spaces need not (cf. Nosofsky, 1991). The Luce choice rule was selected as our generic output model here due to its ubiquity in models of cognitive phenomena—it has been successfully applied to choice behavior ranging from low-level neural networks to high-level economic models of group choice behavior.

The Luce choice rule simulates how humans select from possible choice alternatives given a stimulus similarity space, governed by probabilities conditioned on the choice set. Hence, its input is metric space, but its output is a probability of a given response. Given a set of stimulus similarities (where similarity is defined as an inverse monotonic function of psychological distance) the Luce choice rule states that the probability of responding to stimulus S_i with response R_j is defined as:

$$p(R_j|S_i) = \frac{\beta_j \eta_{i,j}}{\sum_{k \epsilon M} \beta_k \eta_{i,k}} \quad (1)$$

where β_j is the response bias for item j, and $\eta_{i,j}$ is the similarity between stimuli i and j. Given the restrictions of metric spaces, the total probability over all responses sums to one. Most applications of the choice rule include exponential scaling of similarity based on Shepard's (1987) universal law of distance and perceived similarity. Hence, this general formula is often referred to as the Shepard-Luce choice axiom:

$$p(R_j|S_i) = \frac{\beta_j e^{-\lambda d(S_i,R_j)}}{\sum_{k \epsilon M} \beta_k e^{-\lambda d(S_i,R_k)}} \quad (2)$$

[1] Utsumi (2015) has revisited the Steyvers and Tenenbaum (2005) work and demonstrated that while scale-free and small-word structure is unobtainable by LSA, several other variants of the model, all spatial models, naturally produce the correct structure from association norms.

Please cite this article in press as: Jones, M. N., et al., In defense of spatial models of semantic representation, New Ideas in Psychology (2017), http://dx.doi.org/10.1016/j.newideapsych.2017.08.001