# Modeling Word Association Data using Multiple Maps

**Laurens van der Maaten**[1] **Geoffrey Hinton**[2]

Semantic space models such as LSA are not appropriate as computational models for semantic representation because they cannot model (i) intransitive similarities, (ii) objects with high centrality, and (iii) asymmetric similarities.
We present a semantic space model that can represent non-metric similarities by giving each object a location and a weight in multiple maps. The model is implemented as a multiple maps version of t-SNE (a multidimensional scaling variant). We demonstrate the merits of our model on a dataset of word association data.

## Multiple maps t-SNE

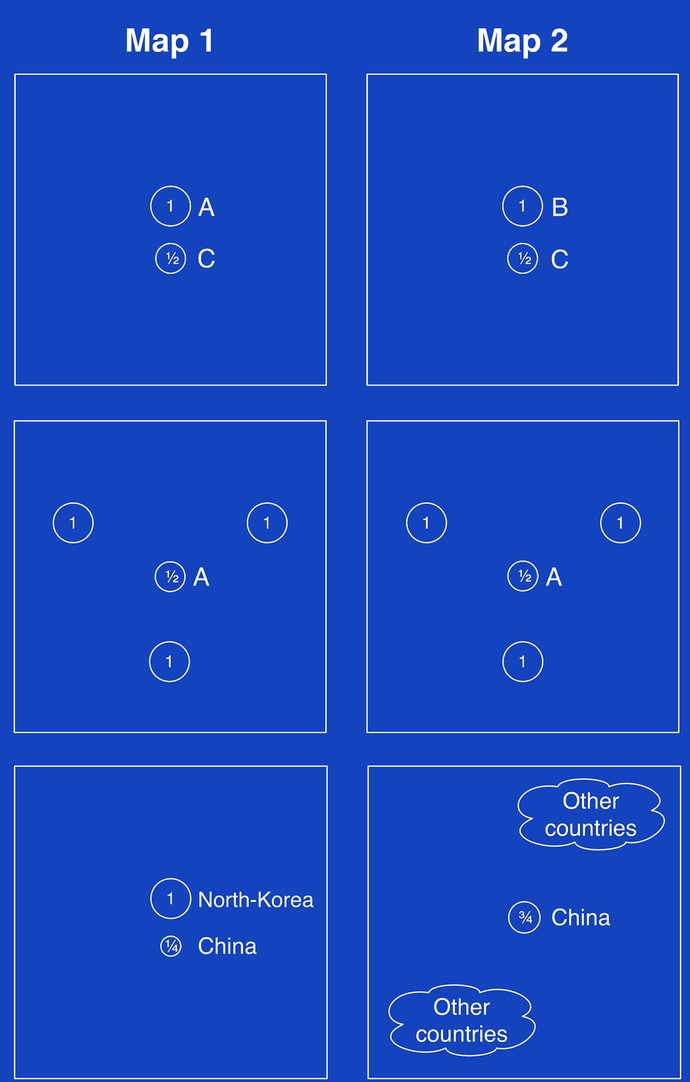
Word association data comprises conditional probabilities $p_{j|i}$ that describe the similarity between objects $i$ and $j$. Under the model, the similarity between $i$ and $j$ is represented by a weighted sum of similarities in the maps:

$$q_{j|i} = \frac{\sum_m \pi_i^{(m)} \pi_j^{(m)} \left(1 + \|y_i^{(m)} - y_j^{(m)}\|^2\right)^{-1}}{\sum_{m'} \sum_{k \neq i} \pi_i^{(m')} \pi_k^{(m')} \left(1 + \|y_i^{(m')} - y_k^{(m')}\|^2\right)^{-1}}$$

Minimize sum of Kullback-Leibler divergences:

$$C = \sum_i KL(P_i || Q_i) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}}$$

## Non-metric similarities




## Word association maps

The model identifies different senses of words. For instance, look up the words *'cheerleader'* or *'monarchy'* in the three maps above.

[1] Tilburg University, The Netherlands. Email: lvdmaaten@gmail.com
[2] University of Toronto, Canada. Email: hinton@cs.toronto.edu
For more information: http://ticc.uvt.nl/~lvdrmaaten/tsne

Tilburg University NWO