Measuring Linguistic Regularity

Erhard Hinrichs

Seminar für Sprachwissenschaft
Eberhard-Karls Universität Tübingen

Computing Analogies

Problem: analogy question a:b :: c:d where d is unknown

Example: apple : tree :: grape : ?

Solution:
Assume vectors x_a, x_b, x_c (all normalized to unit norm), and compute:

$$y = x_b - x_a + x_c \tag{1}$$

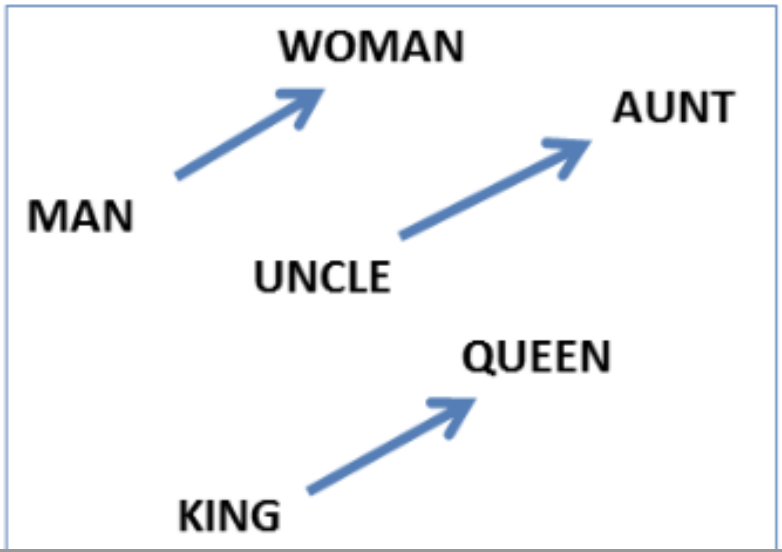
Example:

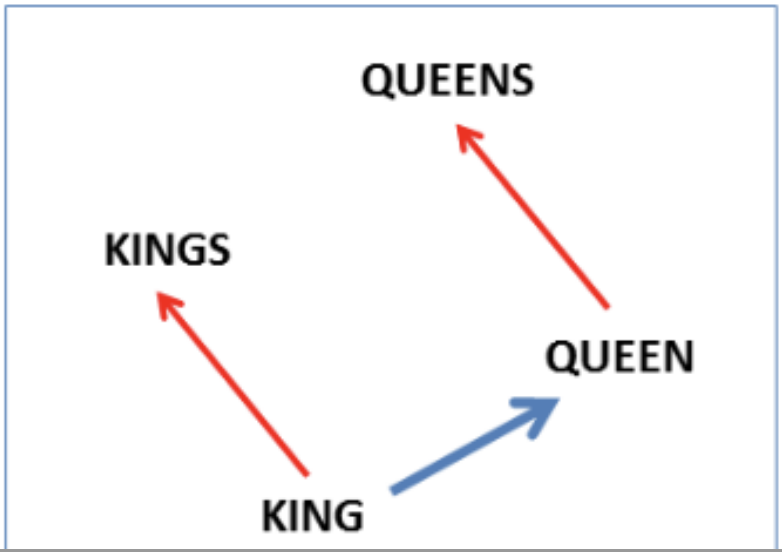
$$y = \mathit{tree} - \mathit{apple} + \mathit{grape} \tag{2}$$

A bit of terminology: Is this really analogy or rather relational similarity?

Observation by Mikolov et al. (2013)

We have found that a simple vector offset method based on cosine distance is remarkably effective in solving these questions. In this method, we assume relationships are present as vector offsets, so that in the embedding space, all pairs of words sharing a particular relation are related by the same constant offset.





Explanation

- y is the continuous space representation of the word we expect to be the best answer under the parallelogram assumption of analogical learning introduced by Rumelhart and Abrahamson (1973).
- Of course, no word might exist at that exact position, so we then search for the word whose embedding vector has the greatest cosine similarity to y and output it.

Argmin or Argmax - that is the question

Mikolov et al. (2013):

$$w^* = argmax_w \frac{x_w y}{|| x_w || || y ||} \tag{3}$$

Jurafsky and Martin (2024):

$$w^* = argmin_x \ distance(x, b - a + a^*) \tag{4}$$

Criticism by Joshua Peterson

- Cosine Similarity obeys the principle of symmetry.
- Human judgments of similarity do not always obey symmetry.
- Informants consider North Korea to be more similar to China, than China to North Korea.

Bibliographical References

Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. 2013. Linguistic Regularities in Continuous Space Word Representations. In *Proceedings of NAACL: Human Language Technologies*, 746–751, Atlanta, Georgia. Association for Computational Linguistics.

Peterson, J. C. , D. Chen, and T. L. Griffiths. 2020. Parallelograms revisited: Exploring the limitations of vector space models for simple analogies. Cognition, 205.

Rumelhart, D. E. and A. A. Abrahamson. 1973. A model for analogical reasoning. Cognitive Psychology, 5(1):1–28..

Tversky, Amos. 1977. Features of similarity. Psychological Review, 84 (4), 327-352