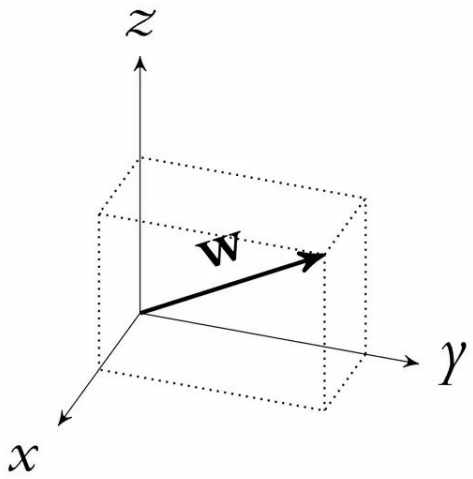
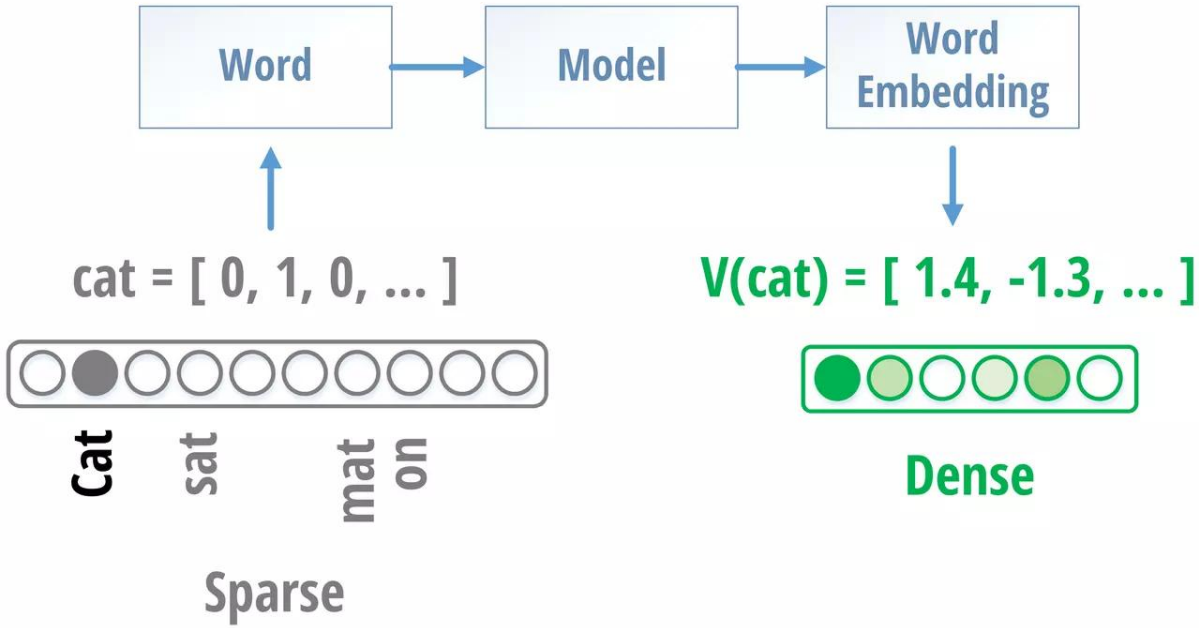
$\vec{W}$ord Embeddings

A non-exhaustive introduction to Word Embeddings



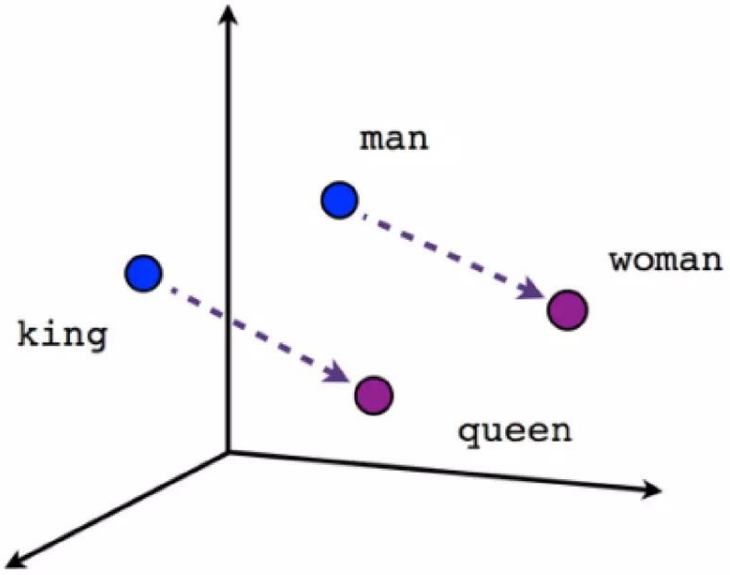
Word Embeddings



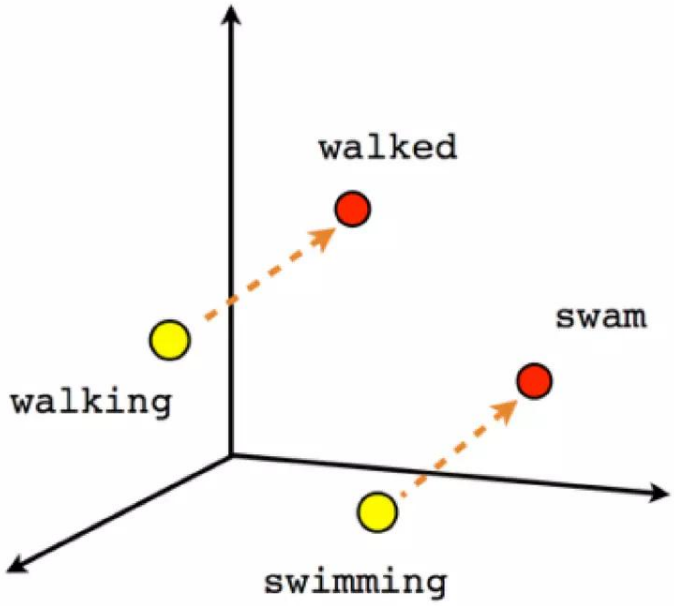
- From a sparse representation (usually one-hot encoding) to a dense representation
- Embeddings created as by-product vs explicit model

Amazing embeddings

Semantic relationships are often preserved on vector operations.



Male-Female



Verb tense

Source: TensorFlow

Word analogies

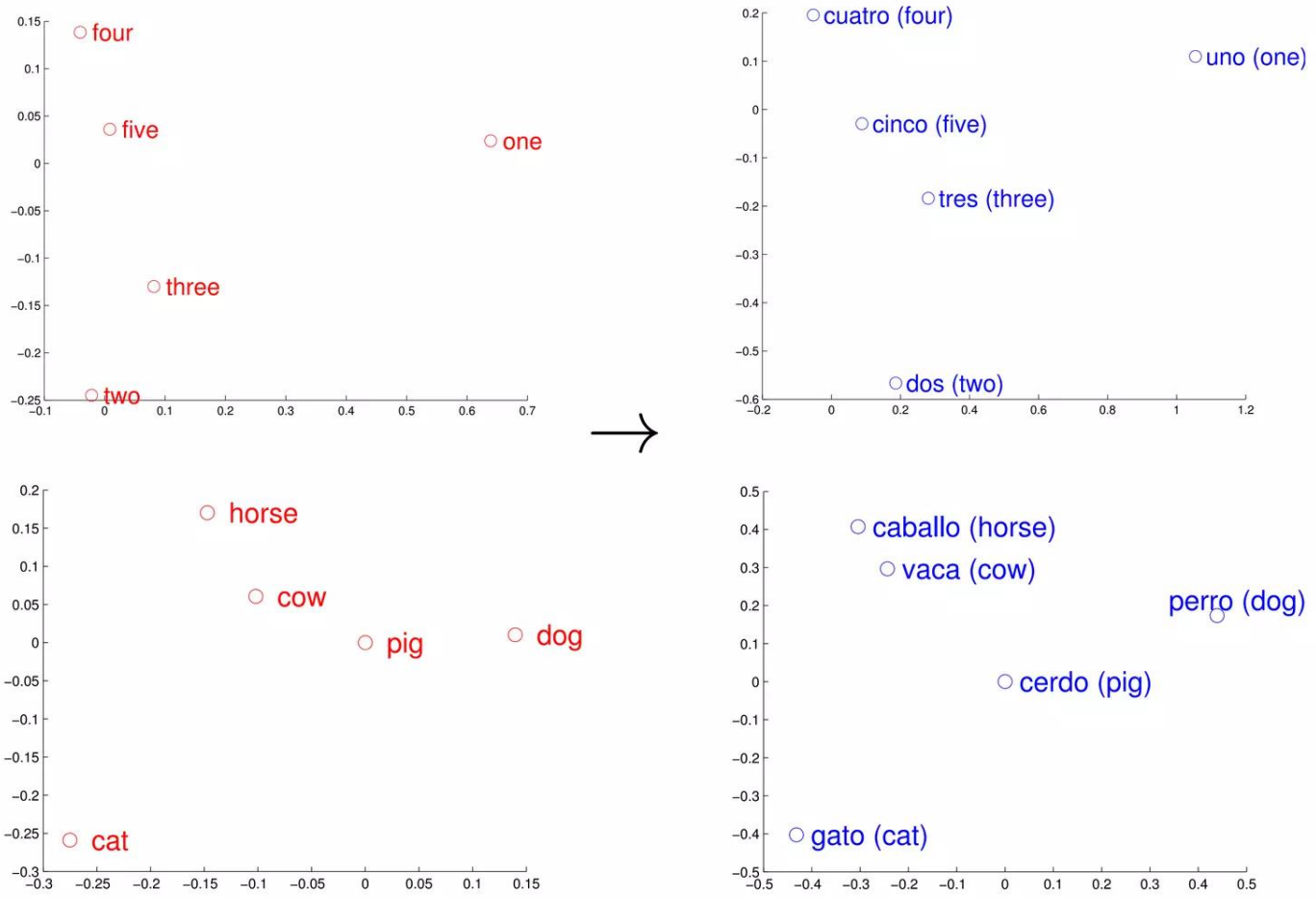
Suppose we have the vector $\vec{w} \in \mathbb{R}^n$ of any given word such as $\vec{w}_{king}$, then we can do:

$$\vec{w}_{king} - \vec{w}_{man} + \vec{w}_{woman} \approx \vec{w}_{queen}$$

This vector operation shows that the closest word vector to the resulting vector is the vector $\vec{w}_{queen}$.

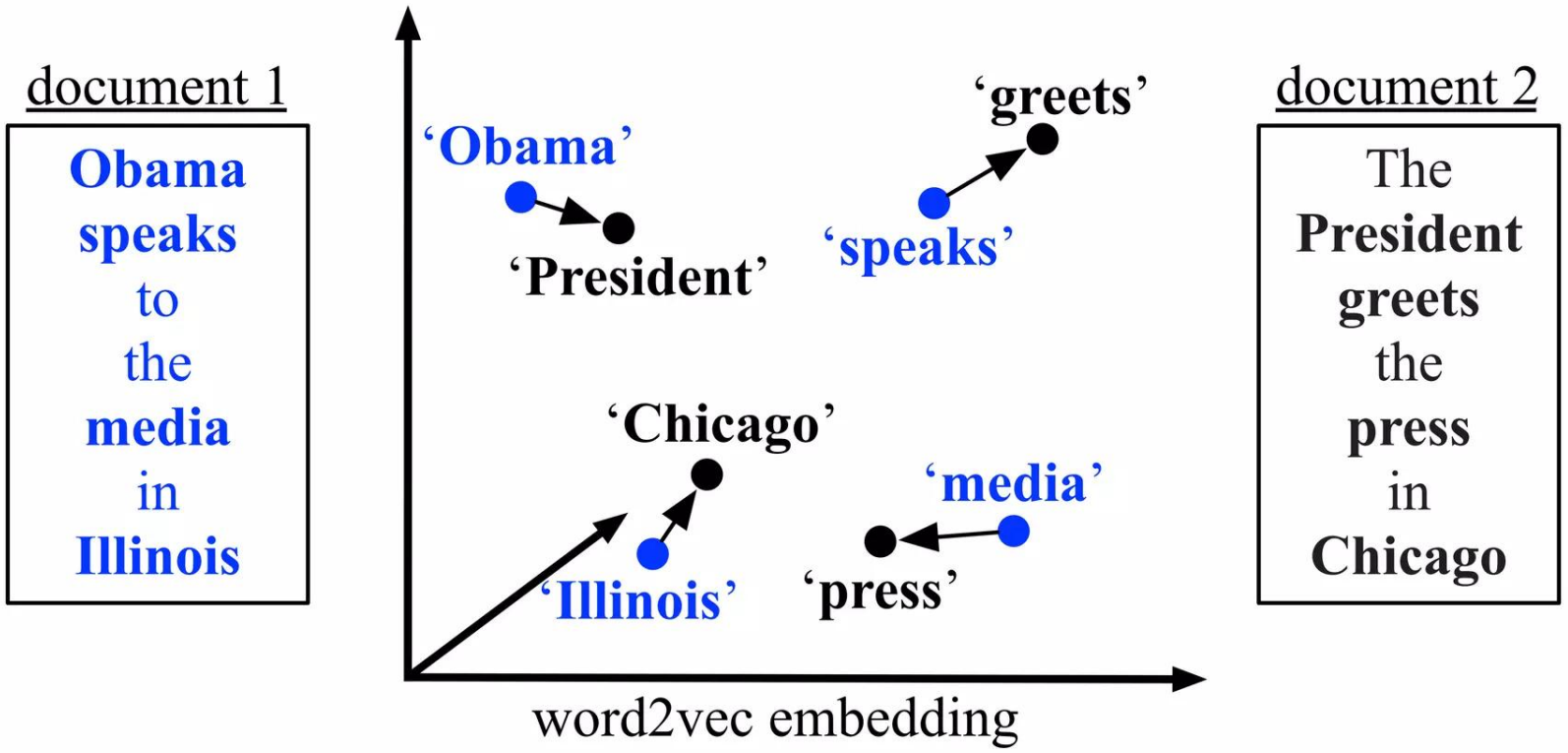
This is an **amazing property** for the word embeddings, because it means that they carry **important relational information** that can be used to many different tasks.

Language structure



Source: Exploiting Similarities among Languages for Machine Translation. Mikolov, Thomas et al. 2013.

Word Movers Distance



Source: From Word Embeddings To Document Distances. Kusner, Matt J. et al. 2015.

Vector Space Model