Word embeddings

Justin Savoie

UofT Political Science Quantitative Methods Research Cluster Meeting

19 November 2021

Plan (part 1) …then… part 2 by Jacob Winter

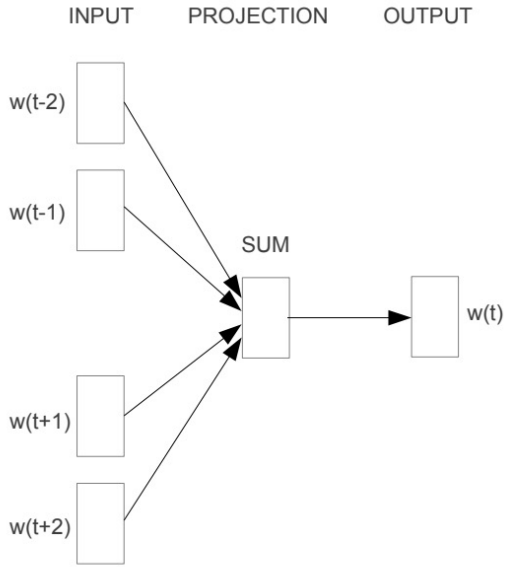
1. Introduction: the method of word embeddings

2. Some theoretical underpinnings: (2.1) linguistic structuralism, (2.2) semantic holism, (2.3) use theory of meaning (2.4)

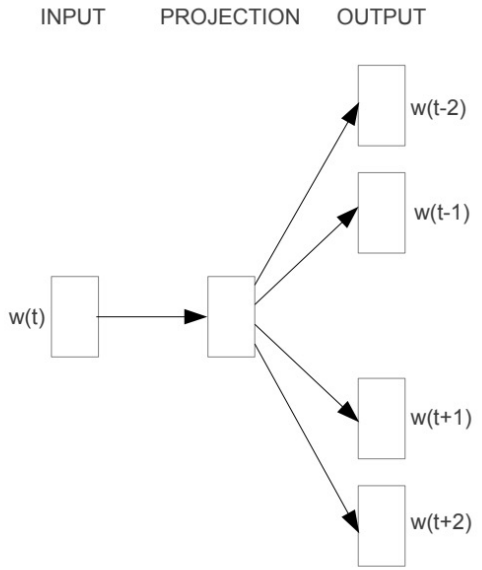
3. Discussion with an example: a model of author style (a comment on *A General Model of Author "Style" with Application to the UK House of Commons, 1935–2018* (Huang, Perry, Spirling, 2020, Political Analysis))

1. Introduction: word embeddings; 4 key points

- **(1)** “quantify and categorize semantic similarities between linguistic items based on their distributional properties in large samples of language” (Firth)
- **(2)** words are vectors in high dimensional space e.g., freedom = [e1,e2,e3,...,e300]
- **(3)** algorithm: predict word or context **(4)** interesting semantic properties

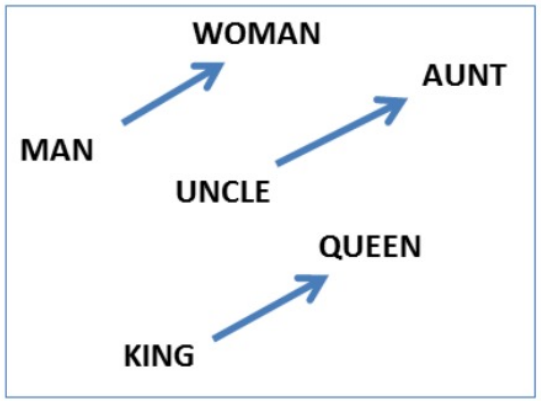


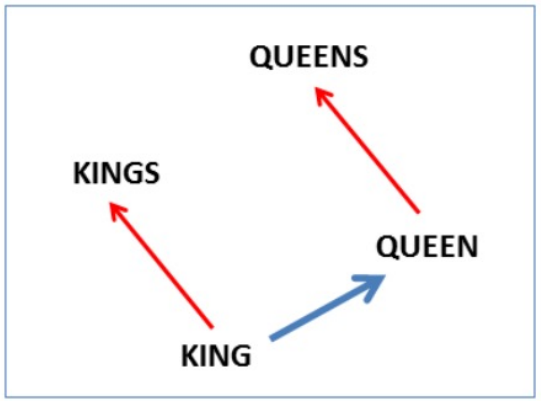
CBOW



Skip-gram

Source: Mikolov, Le, Sutskever (2013)





Source: Mikolov, Yih, Zweig, 2013

1. Introduction: Embeddings in social science

- produce scaling estimates of ideological placement (Rheault, Cochrane, 2020)
- trace the changing cultural meanings of words (Rodman, 2020)
- specify a relational model of meaning consistent with contemporary theories of culture (Kozlowski et al. 2019)

2. Theoretical underpinnings
2.1 Linguistic structuralism

- Arseniev-Koehler (2021, working paper)

- Like any tool, WE build on assumptions. WE operationalize the assumptions from structural linguistic theory of meaning.

- meaning is relational (1) coherent (2) can be analyzed as a static system (3)

- WE offer a theoretical opportunity to think about schema, oppositions, coherence, signs

2.1 Linguistic structuralism

- Linguistic structuralism: language is a system of signs. Language is about relationships between signs. Not about the representation of an external reality.
- 1 meaning is relational
  - WE do learn and represent vectors relationally
- 2 meaning is coherent
  - WE force coherence over contexts (except contextual WE) and over dimensions
- 3 it makes sense to analyze meaning as a static system
  - When freezing and looking at WE, they are static. But training is dynamic.

2.1 Linguistic structuralism: PROBLEMS

- 1 Meaning is grounded or embodied
- 2 Oppositions are reductionistic
- 3 Meaning is not coherence
- 4 Language is dynamic

2.2 Semantic holism

- Musil (2020)

- Meaning is something words have in sentences, is something sentences have in a language.
- Atomism: language = vocabulary + syntax; words in vocabulary have meaning; words have meaning because they ***refer*** to something independent
- Holism accepts interdependence, the whole is the vehicle of meaning

- “The unit of empirical significance is the whole of science.”

2.2 Semantic holism

- Musil’s specific point:
  - The distributional hypothesis works, but we don’t really know why
  - Holism is great when there is nothing beside language to anchor meaning to
  - The truth-value potential in Frege’s sentence holism: two expressions have the same truth-value potential (meaning) iff whenever each is completed by the same expression to form a sentence, the two sentences have the same truth-value **corresponds to Skip-gram**, not CBOW
  - Skip-gram fixes a word and predicts context, so it learns the meaning of the fixed word
  - CBOW fixes the context and predicts word, so CBOW learns the meaning of the context, which isn’t what we want
  - Well known result that CBOW performs worse, especially on rare words, but quicker to fit.
  - When I fix the word, and I ”circulate” contexts, I learn which contexts fit better. I learn the meaning. By fixing the context, we do the opposite.

2.3 The use theory of meaning

- The meaning of a word is it use

- ‘To say “This combination of words makes no sense” excludes it from the sphere of language and thereby bounds the domain of language. [. . .] When a sentence is called senseless, it is not as it were its sense that is senseless. But a combination of words is being excluded from the language, withdrawn from circulation.’ PI 499-500

2.3 The use theory of meaning

- Skelac & Jandric (2020)
- “although all of them [word2vec, Firth, Wittgenstein] emphasise the importance of context, its scope is differently understood.”

- With Firth: “Word2vec offers a restricted view of what constitutes a context … it is limited to directly neighbouring words only”
- With Wittgenstein: language games are not about collocation

3. Discussion

- Research in Poli Sci; levels:
- (1) content of meaning (2) structure of meaning (oppositions, vs scales) (3) coherence
- (2) and (3) can be inspired by theoretical work like this

- Example: A model of author style (Huang et al.)
  - Intuition: Is it easy or hard to predict the author of text? If it's easy, the author is distinctive
- Possible examples:
  - comparing word2vec with contextualized embeddings: "richness of vocabulary"
  - multimodal embeddings

Work cited

- Alina Arseniev-Koehler (2021) Theoretical foundations and limits of word embeddings: what types of meaning can they capture? Working paper https://arxiv.org/pdf/2107.10413.pdf

- Timo Honkela (2007) Philosophical Aspects of Neural, Probabilistic and Fuzzy Modeling of Language Use and Translation https://ieeexplore.ieee.org/abstract/document/4371417?casa_token=MXw67lLQOLUAAAAA:SyOr2K-ecDWQsx1YNGqngTPkr103oIVHayEqBqa61yxa9RH64XUZREsIAWlfhDRDIyDQpM_5oO8

- Tomáš Musil (2020) Semantic Holism and Word Representations in Artificial Neural Networks https://arxiv.org/pdf/2003.05522.pdf

- Skelac, I., & Jandrić, A. (2020). Meaning as Use: From Wittgenstein to Google's Word2vec. In *Guide to Deep Learning Basics* (pp. 41-53). Springer, Cham.

Resources

- Using Word Embeddings in Political Science http://arthurspirling.org/documents/embed.pdf
- Using Python's gensim https://radimrehurek.com/gensim/auto_examples/tutorials/run_word2vec.html#sphx-glr-auto-examples-tutorials-run-word2vec-py
- Lots! of great tutorials on github (can find on google) here is one by Arseniev-K discussed earlier https://github.com/arsena-k/Exploring_WordEmbeddings/blob/master/Exploring%20Word%20Embeddings%20and%20Polysemy.ipynb