Dynamic and Static Prototype Vectors for Semantic Composition

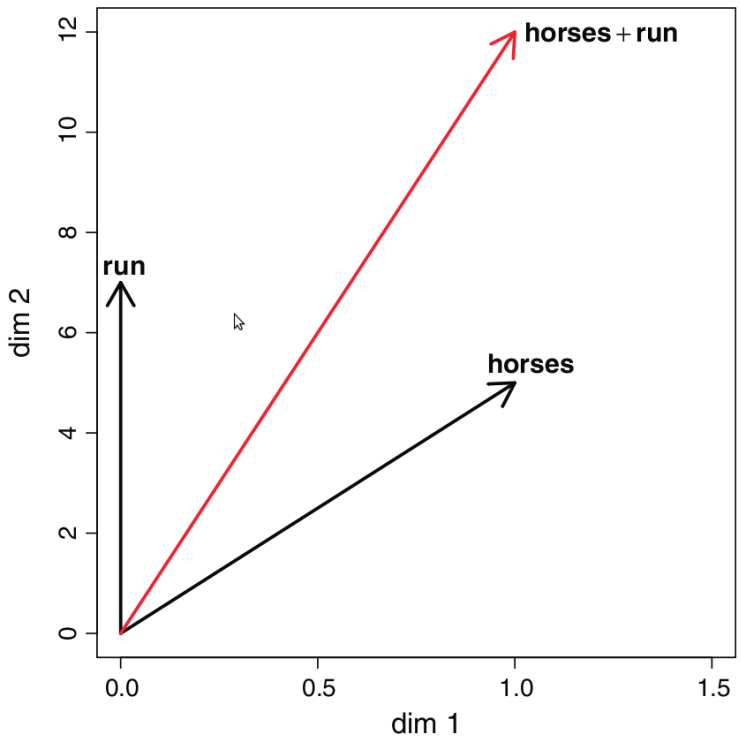
by S. Reddy, I. P. Klapaftis, D. McCarthy and S. Manandhar (2011)

presentation at CL reading group by S. Suster
22/11/2013

Context

A word is a vector. . .

Composition is a combination of word vectors



Idea

Problem: single vector conflates all senses, but for composition, only some are relevant

- Can we somehow pick the right sense?
  - [Words mutually determine each other's senses (co-compositionality)]
- Would this lead to better composition?

Exemplars. Prototypes.

Exemplar: vector representing one context of a particular word in the corpus

Exemplar-based modeling:

1. keeping exemplars
2. selectively choosing exemplars before combining them (e.g. addressing polysemy by removing irrelevant exemplars)

Prototype vector: vector generalized from exemplars

Method

1. Represent a sense with a prototype vector
2. Compose (addition and multiplication)

Prototypes obtained in 2 ways:

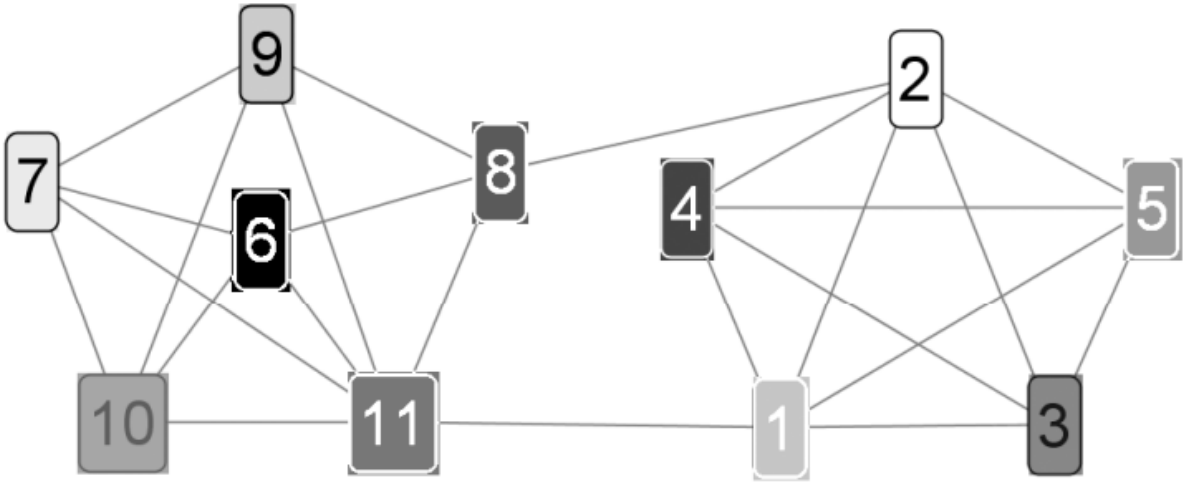
- Static:
  - Have many prototypes for some noun and then choose correct ones
  - Prototypes obtained once, their number is fixed for a particular noun
- Dynamic: Build one (correct) prototype on the fly

Static prototypes

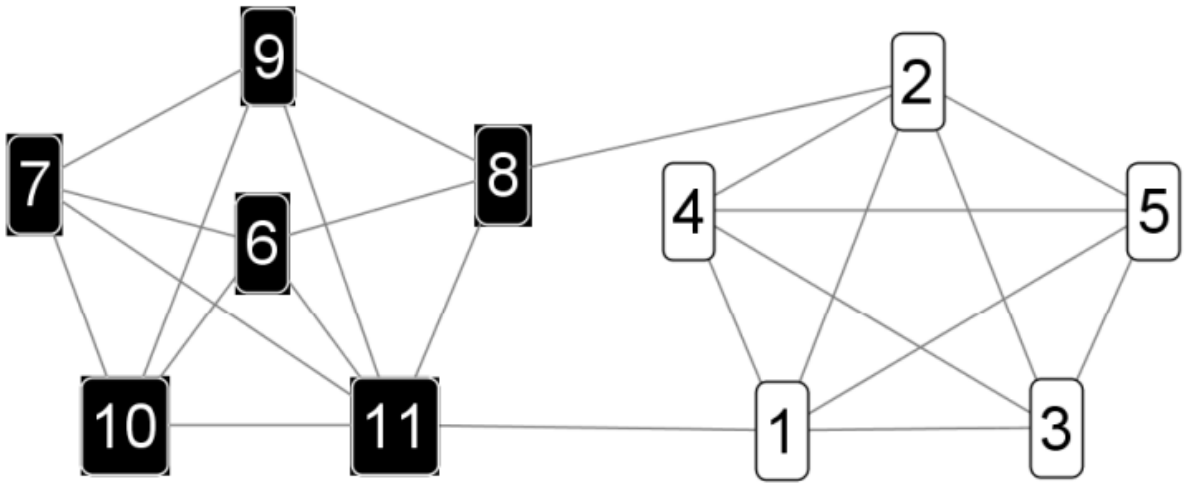
Induce word senses for a noun:

1. build graph with contexts as nodes:
   - highly similar contexts $\Rightarrow$ connected, with high weight
   - weight determined by both collocation and word overlap
2. clustering on the obtained graph
   - (# of clusters unknown in advance $\Rightarrow$ choice of clustering algorithm)
   - a node inherits the class of that adjacent node which has the highest sum of edge weights

11 contexts, 11 classes



11 contexts, 2 classes



Representing classes. Relevant sense.

- For each class, a prototype is determined
- This is just the centroid vector (mean)

- What are the right classes (senses) for noun compound?
- Senses that are the most similar (cosine)

Dynamic prototypes I

- No fixed set of senses for a word $\Rightarrow$ every context yields a slightly different sense
- Sense is built with the help of other constituent:
  1. Refine the set of exemplars of one noun based on the other noun
  2. Take centroid to obtain the prototype

Dynamic prototypes II

Refinement is achieved by ranking exemplars based on overlap/similarity:

- between a) exemplar of the first constituent and b) collocates of the second constituent
- between a) exemplar of the first constituent and b) words similar to the second consituent

- Choose top n% of ranked exemplars to build a prototype
- Same procedure for the second constituent

Evaluation

- Dataset with human ratings for similarity between pairs of compound nouns
- Model is evaluated by:
  - calculating similarity between composed vectors
  - correlating with human scores

Baselines:

- 1 vector/word: conflates all senses
- 1 vector/compound (no composition)
  - if this should give the best performance, why do we need composition?

Outcome

- Costructing dynamic prototypes is simpler than static prototypes
- Yet dynamic prototypes are much better

Poor quality of static prototypes because of

- sense selection?
- cluster granularity?
- verbs adding to noise?
- x?

Reflection. Discussion.

Positive:

- Overall well written
- Nice ideas
- Contrasting static and dynamic sense construction
- “Dynamic prototypes” is very intuitive and well performing!

Reflection. Discussion.

Following could be further improved:

- Comparison to other strong performing methods
- Extrinsic evaluation
- Beyond compound nouns? (in static prototypes, the correct senses of the two nouns are chosen based on their similarity)
- Amount of contribution of one constituent towards the compound
- Is the composition function powerful enough?
- Some compound nouns can hardly be disambiguated without context:

“Mouse game”

Which sense of mouse/game?