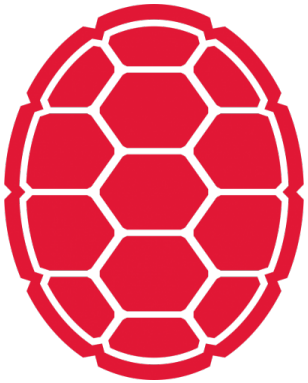


Distributional Semantics

Machine Learning: Jordan Boyd-Graber
University of Maryland
SLIDES ADAPTED FROM YOAV GOLDBERG AND OMER LEVY

Beyond word2vec

- word2vec is factorizing a word-context matrix.
- The content of this matrix affects the resulting similarities.
- word2vec allows you to specify a *window size*.
- But what about other types of contexts?

- Example: **dependency contexts** (Levy and Dagan, ACL 2014)

Bag of Words (BoW) Context

Australian scientist discovers star with telescope

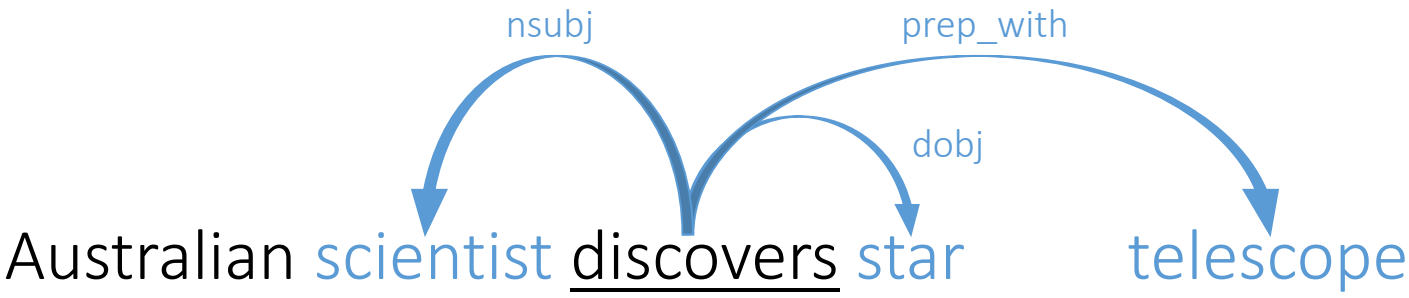
Bag of Words (BoW) Context

Australian scientist discovers star with telescope

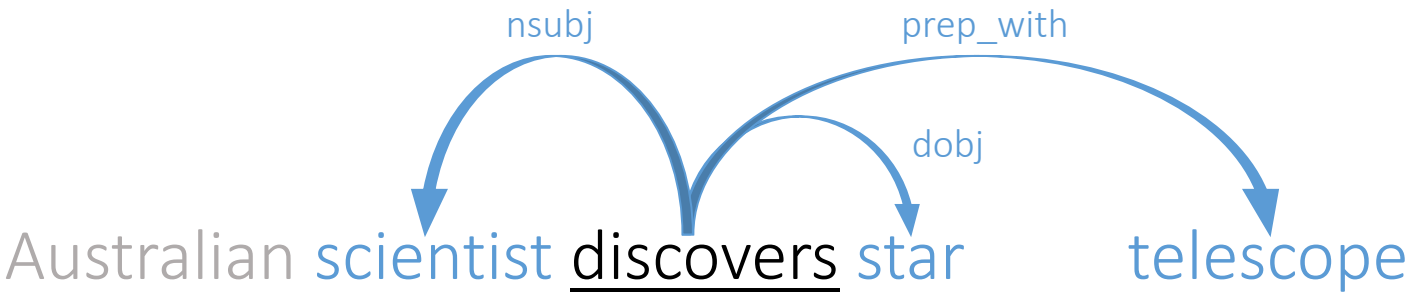
Syntactic Dependency Context

Australian scientist discovers star with telescope

Syntactic Dependency Context



Syntactic Dependency Context



Embedding Similarity with Different Contexts

Target Word	Bag of Words (k=5)	Dependencies
Hogwarts (Harry Potter's school)	Dumbledore hallows half-blood Malfoy Snape	Sunnydale Collinwood Calarts Greendale Millfield
	Related to Harry Potter	**Schools**

Embedding Similarity with Different Contexts

Target Word	Bag of Words (k=5)	Dependencies
Turing (computer scientist)	nondeterministic non-deterministic computability deterministic finite-state	Pauling Hotelling Heting Lessing Hamming
	Related to computability	**Scientists**

Embedding Similarity with Different Contexts

Target Word	Bag of Words (k=5)	Dependencies
dancing (dance gerund)	singing dance dances dancers tap-dancing	singing rapping breakdancing miming busking
	Related to dance	**Gerunds**

Online Demo!

Context matters

Choose the correct contexts for your application

- larger window sizes – more topical
- dependency relations – more functional

Context matters

Choose the correct contexts for your application

- larger window sizes – more topical
- dependency relations – more functional
- only noun-adjective relations
- only verb-subject relations

Context matters

Choose the correct contexts for your application

- larger window sizes – more topical
- dependency relations – more functional
- only noun-adjective relations
- only verb-subject relations
- context: time of the current message
- context: user who wrote the message

Context matters

Choose the correct contexts for your application

- larger window sizes – more topical
- dependency relations – more functional
- only noun-adjective relations
- only verb-subject relations
- context: time of the current message
- context: user who wrote the message
- . . .
- the sky is the limit

Summary

Distributional Semantics

- Words in similar contexts have similar meanings.
- Represent a word by the contexts it appears in.
- But what is a context?

Neural Models (word2vec)

- Represent each word as dense, low-dimensional vector.
- Same intuitions as in distributional vector-space models.
- Efficient to run, scales well, modest memory requirement.
- Dense vectors are convenient to work with.
- Still helpful to think of the context types.