Symmetric Patterns *and* Coordinations: Fast *and* Enhanced Representations of Verbs *and* Adjectives

Roy Schwartz*, Roi Reichart+ and Ari Rappoport*, ***NAACL 2016***

*The Hebrew University, +Technion

Contribution

- Training **word2vec** with **symmetric pattern** contexts improves verb similarity performance by ***15%***
- Training with *symmetric pattern* contexts is also **30-50 times** faster than other context types
- Fully unsupervised *symmetric pattern* contexts are even better than (supervised) syntactic coordination contexts

Background and Motivation

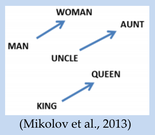
Word Embeddings are Great!

Most similar words to ***france***:

Word	distance
spain	0.678515
belgium	0.665923
netherlands	0.652428
italy	0.633130
switzerland	0.622323
luxembourg	0.610033
portugal	0.577154
russia	0.571507

(Mikolov et al., 2013)

(Turian et al., 2010)



(Mikolov et al., 2013)

But… what about Verbs?

- State-of-the-art word embeddings perform poorly on verb similarity
- On SimLex999 (Hill et al., 2015):

verbs << nouns

Model	Nouns	Verbs
GloVe	0.38	**0.16**
word2vec skip-gram	0.50	**0.31**
SP+ (**Schwartz** et al., 2015)	0.50	0.58

Other models do much better

Context Type in Word Embeddings

- Word embeddings models are trained on (*word, context*) pairs
- Leading embedding models train using **bag-of-words** contexts
- Other options exist
  - *Dependency links* (Levy & Goldberg, 2014)
  - *Symmetric patterns* (**Schwartz** et al., 2015)
- We study the effect of the context type on verb similarity performance

Symmetric Patterns

Davidov and Rappoport (2006)

- Words that co-occur in symmetric patterns often take the same semantic role
  - *John **and** Mary* went to school
  - Is it better to *walk **or** run*?
  - Jane is *smart **as well as** funny*
- Symmetric patterns also capture different aspects of word **similarity** (Davidov & Rappoport; 2006, Feng et al., 2013; **Schwartz** et al., 2014;2015)

X *and* **Y**

beds *and* **sofas**

sofas *and* **beds**

Experiments

- Experiments with the skip-gram model (Mikolov et al., 2013)
  - Each time with a different context type
  - All other modeling decisions **are identical**
- Models train on an 8G words corpus
- Experiments with the verb portion of SimLex999

Symmetric Patterns vs. Coordinations

- Symmetric patterns are an **unsupervised** estimation of syntactic coordinations
- Coordinations are more effective than all dependency links on verbs and adjectives
- Even though coordinations are captured using a **supervised** parser, Symmetric patterns are **more useful** for verb and adjective similarity (and more compact!)

Additive value of Context Type and Antonym Detection

- In **Schwartz** et al. (2015) we have shown superior verb similarity results using symmetric pattern contexts, but also using an **antonym** detection mechanism
- We re-implemented that model without the antonym feature (SP-)
- *Symmetric patterns* and the *antonym* detection feature have an **additive value**

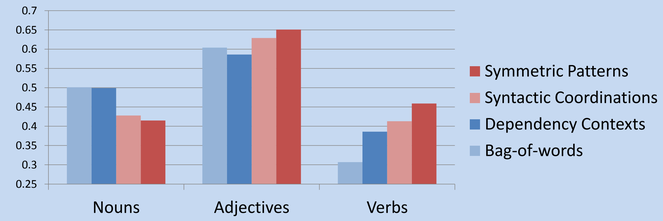
Context Type	Model	Spearman's ρ	Delta
Bag-of-Words	skip-gram	0.31	-
Symmetric Patterns	skip-gram	0.46	15%
	SP- (Schwartz et al., 2015)	0.44	13%
Symmetric Patterns + Antonyms	SP+ (Schwartz et al., 2015)	**0.58**	**27%**

Results and Discussion

Context Type Matters!

Symmetric Patterns >> Bag-of-words

- Up to 15% improvement on verbs and 9% on adjectives



Compact Model

Symmetric Pattern Contexts are *Super Fast* to Train

- 30-50 times faster than bag-of-words and dependency links

Context Type	Num of Contexts	Train Time (Mins)
Bag-of-words	13B	320
Dependency links	14.5B	551
Syntactic Coordinations	0.55B	23
Symmetric Patterns	**0.27B**	**11**

Main Take Home Message

- The **context** type plays **an important role** in word similarity: Symmetric pattern contexts yield **15%** verb similarity improvement and **9%** adjective similarity improvement
- Symmetric pattern train **much** faster than other context types (only 2-3% of the training time)
- Symmetric patterns are even better than (**supervised**!) coordination structures

Future Work

- Automatic combination of context types for improved word similarity (Vulić et al., in review)
- What is a pattern? Empirical evaluation of pattern types
- www.cs.huji.ac.il/~roys02