Vector Space Models of Lexical Meaning - Part II

Lukas Schmitt

Saarland University

s9lkscmi@stud.uni-saarland.de

November 20, 2017

Overview

1. Compositional Vector Space Models
2. Background
3. The Compositional Framework
4. Conclusion

Compositional Vector Space Models

Development in Computational Linguistics

Compositional treatment for distributional vector-based models

Compositional Vector Space Models

Development in Computational Linguistics

Compositional treatment for distributional vector-based models

Framework by Coecke et al.

Designed for linguists and has interesting links with Montague semantics

Compositional Vector Space Models

Development in Computational Linguistics

Compositional treatment for distributional vector-based models

Framework by Coecke et al.

Designed for linguists and has interesting links with Montague semantics

Goal of the Framework

Get one vector for a whole sequence of words based on the vectors of each word

Compositional Vector Space Models

Development in Computational Linguistics

Compositional treatment for distributional vector-based models

Framework by Coecke et al.

Designed for linguists and has interesting links with Montague semantics

Goal of the Framework

Get one vector for a whole sequence of words based on the vectors of each word

But why?

Compositional Vector Space Models

Reasons

- Interesting new viewpoint on the problem
- Language Processing applications benefit from it

Compositional Vector Space Models

Reasons

- Interesting new viewpoint on the problem
- Language Processing applications benefit from it

$\Downarrow$

Query: Find all car showrooms with sales on for Ford cars.

Compositional Vector Space Models

Reasons

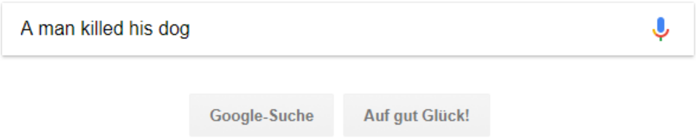
- Interesting new viewpoint on the problem
- Language Processing applications benefit from it

$$\Downarrow$$

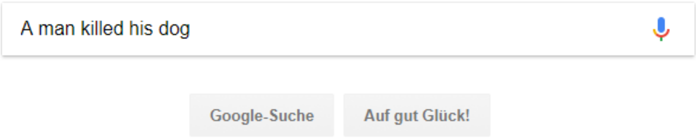
Query: Find all car showrooms with sales on for Ford cars.

Headline: Cheap Fords available at the car salesroom.

Compositional Vector Space Models

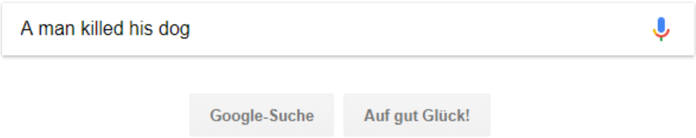


Compositional Vector Space Models



$$\Downarrow$$

Compositional Vector Space Models

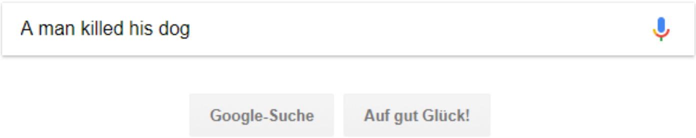


$$\Downarrow$$

Top-Ranked page: Dog shoots man.

Third-Ranked page: The Man who killed his friend for eating his dog after it was killed.

Compositional Vector Space Models



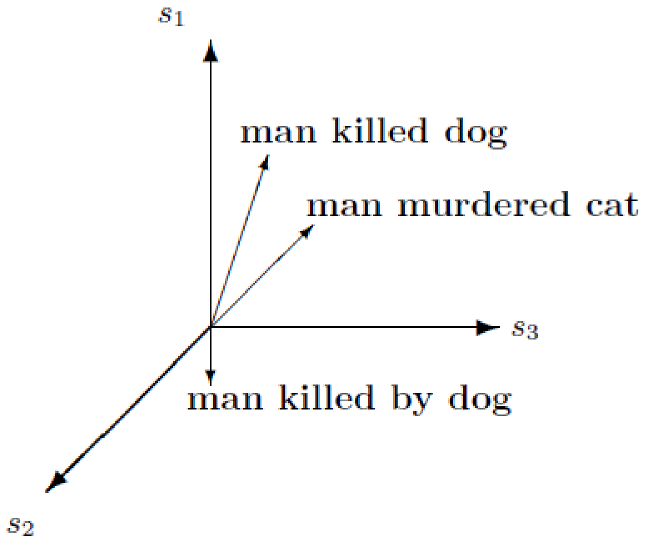
$\Downarrow$

Top-Ranked page: Dog shoots man.

Third-Ranked page: The Man who killed his friend for eating his dog after it was killed.

$\rightarrow$ Order of words matter!

Compositional Vector Space Models



Distributional Sentence Representation

Intuition

Last talk: Words that occur in similar contexts tend to have similar meanings

Distributional Sentence Representation

Intuition

Last talk: Words that occur in similar contexts tend to have similar meanings

$\Downarrow$

How to apply to a larger sequences of words?

Existing Vector Composition Methods

Information Retrieval

Vector addition is used for combination of word vectors, but also other binary vector operations

Existing Vector Composition Methods

Information Retrieval

Vector addition is used for combination of word vectors, but also other binary vector operations

Cognitive Science

How can distributional, connectionist representations be given a compositional treatment to reflect the compositional nature of language?

$$\downarrow$$

Tensor Product

Existing Vector Composition Methods

Information Retrieval

Vector addition is used for combination of word vectors, but also other binary vector operations

Cognitive Science

How can distributional, connectionist representations be given a compositional treatment to reflect the compositional nature of language?

$\downarrow$

Tensor Product

Baroni & Zamparelli

Usage of matrix multiplication of vectors and machine learning methods

The Compositional Framework

syntax drives the compositional process

The Compositional Framework

syntax drives the compositional process

Syntactic and semantic descriptions are *type-driven*

The Compositional Framework

syntax drives the compositional process

Syntactic and semantic descriptions are *type-driven*

Syntactic Type in Combinatory Categorial Grammar

$likes := (S \backslash NP)/NP$

The Compositional Framework

syntax drives the compositional process

Syntactic and semantic descriptions are *type-driven*

Syntactic Type in Combinatory Categorial Grammar

$likes := (S\backslash NP)/NP$

Classical Categorial Grammar

$X/Y \; Y \Rightarrow X \; (>)$
$Y \; X\backslash Y \Rightarrow X \; (<)$

The Compositional Framework

Adjoint Categories

$likes := NP^r \cdot S \cdot NP^l$

The Compositional Framework

Adjoint Categories

$likes := NP^r \cdot S \cdot NP^l$

Cancellation Rules

$X \cdot X^r \to 1$
$X^l \cdot X \to 1$

Pregroup Derivation

$$
\begin{array}{ccccccc}
\text{Investors} & \text{are} & \text{appealing} & \text{to} & \text{the} & \text{Exchange} & \text{Commission} \\
\hline
NP & NP^r \cdot S[dcl] \cdot S[ng]^l \cdot NP & NP^r \cdot S[ng] \cdot PP^l & PP \cdot NP^l & NP \cdot N^l & N \cdot N^l & N \\
\cline{6-7}
 & & & & & \multicolumn{2}{c}{N} \\
\cline{5-7}
 & & & & \multicolumn{3}{c}{NP} \\
\cline{4-7}
 & & & \multicolumn{4}{c}{PP} \\
\cline{3-6}
 & & \multicolumn{4}{c}{NP^r \cdot S[ng]} & \\
\cline{2-5}
 & \multicolumn{4}{c}{NP^r \cdot S[dcl]} & & \\
\cline{1-4}
\multicolumn{4}{c}{S[dcl]} & & & \\
\end{array}
$$

Semantic Types and Tensor Products

Remember Syntactic Type

$likes := NP^r \cdot S \cdot NP^l$

Semantic Types and Tensor Products

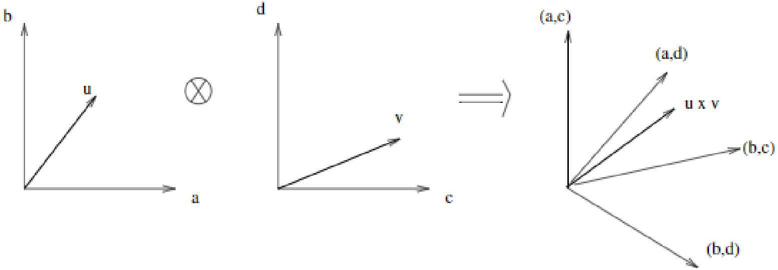
Remember Syntactic Type

$likes := NP^r \cdot S \cdot NP^l$

Vector Space

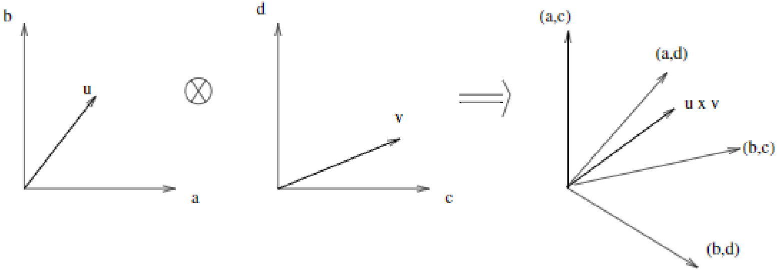
$\overrightarrow{likes} \in \mathbf{N} \cdot \mathbf{S} \cdot \mathbf{N}$

Tensor Product



$$(u \otimes v)_{(a,d)} = u_a \, . \, v_d$$

Tensor Product



$$(u \otimes v)_{(a,d)} = u_a \cdot v_d$$

$$\Rightarrow \overrightarrow{likes} \in \mathbf{N} \otimes \mathbf{S} \otimes \mathbf{N}$$

Individual Vector

$$\Psi = \sum_{ijk} C_{ijk}(\overrightarrow{n_i} \otimes \overrightarrow{s_j} \otimes \overrightarrow{n_k}) \in \mathbf{N} \otimes \mathbf{S} \otimes \mathbf{N}$$

Individual Vector

$$\Psi = \sum_{ijk} C_{ijk}(\overrightarrow{n_i} \otimes \overrightarrow{s_j} \otimes \overrightarrow{n_k}) \in \mathbf{N} \otimes \mathbf{S} \otimes \mathbf{N}$$

$$\Downarrow$$

Answer to the question: What is the semantic type corresponding to a particular syntactic type?

Individual Vector

$$\Psi = \sum_{ijk} C_{ijk}(\overrightarrow{n_i} \otimes \overrightarrow{s_j} \otimes \overrightarrow{n_k}) \in \mathbf{N} \otimes \mathbf{S} \otimes \mathbf{N}$$

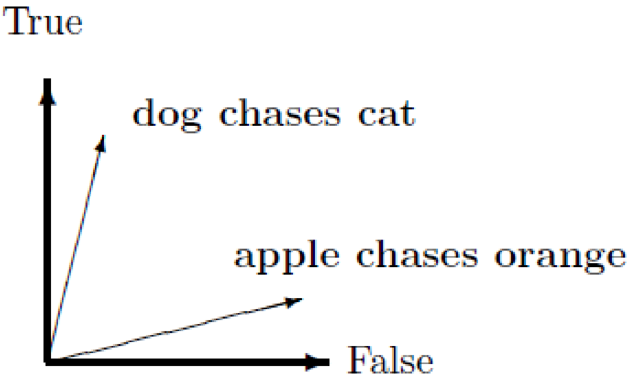
$$\Downarrow$$

Answer to the question: What is the semantic type corresponding to a particular syntactic type?

$$\Downarrow$$

Remaining question: How can the vector for the transitive verb shown above be combined with vectors of a subject and object to give a vector in the sentence space?

Plausability Space for Sentences



Noun Vectors in N

	fluffy	run	fast	aggressive	tasty	buy	juice	fruit
$\overrightarrow{dog}$	0.8	0.8	0.7	0.6	0.1	0.5	0.0	0.0
$\overrightarrow{cat}$	0.9	0.8	0.6	0.3	0.0	0.5	0.0	0.0
$\overrightarrow{apple}$	0.0	0.0	0.0	0.0	0.9	0.9	0.8	1.0
$\overrightarrow{orange}$	0.0	0.0	0.0	0.0	1.0	0.9	1.0	1.0

Transitive Verb Vectors

	⟨fluffy,T,fluffy⟩	⟨fluffy,F,fluffy⟩	⟨fluffy,T,fast⟩	⟨fluffy,F,fast⟩	⟨fluffy,T,juice⟩	⟨fluffy,F,juice⟩	⟨tasty,T,juice⟩ ...
$\overrightarrow{chases}$	0.8	0.2	0.75	0.25	0.2	0.8	0.1
$\overrightarrow{eats}$	0.7	0.3	0.6	0.4	0.9	0.1	0.1

	⟨fluffy,T,fluffy⟩	⟨fluffy,F,fluffy⟩	⟨fluffy,T,fast⟩	⟨fluffy,F,fast⟩	⟨fluffy,T,juice⟩	⟨fluffy,F,juice⟩	⟨tasty,T,juice⟩ ...
$\overrightarrow{chases}$	0.8	0.2	0.75	0.25	0.2	0.8	0.1
dog, cat	0.8,0.9	0.8,0.9	0.8,0.6	0.8,0.6	0.8,0.0	0.8,0.0	0.1,0.0

Calculation of Basis Vectors

True Basis Vector

$$\overrightarrow{dog\ chases\ cat}_{True} = 0.8.0.8.0.9{+}0.75.0.8.0.6{+}0.2.0.8.0.0{+}0.1.0.1.0.0{+}\ldots$$

False Basis Vector

$$\overrightarrow{dog\ chases\ cat}_{False} = 0.2.0.8.0.9 + 0.25.0.8.0.6 + 0.8.0.8.0.0 + \ldots$$

Combining a Transitive Verb with its Subject and Object

$$
\begin{aligned}
f\left(\overrightarrow{\pi} \otimes \overrightarrow{\Psi} \otimes \overrightarrow{o}\right) &= \sum_{ijk} C_{ijk} \langle \overrightarrow{\pi} \mid \overrightarrow{\pi_i} \rangle \, \overrightarrow{s_j} \, \langle \overrightarrow{o} \mid \overrightarrow{o_k} \rangle \\
&= \sum_{j} \left(\sum_{ik} C_{ijk} \langle \overrightarrow{\pi} \mid \overrightarrow{\pi_i} \rangle \langle \overrightarrow{o} \mid \overrightarrow{o_k} \rangle \right)
\end{aligned}
$$

Pregroup Derivation with semantics

$$
\cfrac{\cfrac{man}{\begin{matrix} NP \\ \mathbf{N} \end{matrix}} \quad \cfrac{\cfrac{bites}{\begin{matrix} NP^r \cdot S \cdot NP^l \\ \mathbf{N} \otimes \mathbf{S} \otimes \mathbf{N} \end{matrix}} \quad \cfrac{dog}{\begin{matrix} NP \\ \mathbf{N} \end{matrix}}}{\begin{matrix} NP^r \cdot S \\ \mathbf{N} \otimes \mathbf{S} \end{matrix}}}{\begin{matrix} S \\ \mathbf{S} \end{matrix}}
$$

The Meaning of a Sentence

$$\overrightarrow{w_1 \cdots w_n} = F(\alpha)(\overrightarrow{w_1} \otimes \cdots \otimes \overrightarrow{w_n})$$

Conclusion

Conclusion

- Active area of research

Conclusion

- Active area of research
- Provide a level of robustness that logic-based approaches can't provide

Conclusion

- Active area of research
- Provide a level of robustness that logic-based approaches can't provide
- Classical logic-based enterprise in NLP has failed

Conclusion

- Active area of research
- Provide a level of robustness that logic-based approaches can't provide
- Classical logic-based enterprise in NLP has failed
- Open question: Whether fundamental concepts from semantics can be incorporated

Conclusion

- Active area of research
- Provide a level of robustness that logic-based approaches can't provide
- Classical logic-based enterprise in NLP has failed
- Open question: Whether fundamental concepts from semantics can be incorporated
- Incorporation of vision into the vector representations

Conclusion

- Active area of research
- Provide a level of robustness that logic-based approaches can't provide
- Classical logic-based enterprise in NLP has failed
- Open question: Whether fundamental concepts from semantics can be incorporated
- Incorporation of vision into the vector representations
- A semantic theory incorporating distributional representations at the word, phrase, sentence or even document level, multi-modal features and effective robust methods of inference

Conclusion

- Active area of research
- Provide a level of robustness that logic-based approaches can't provide
- Classical logic-based enterprise in NLP has failed
- Open question: Whether fundamental concepts from semantics can be incorporated
- Incorporation of vision into the vector representations
- A semantic theory incorporating distributional representations at the word, phrase, sentence or even document level, multi-modal features and effective robust methods of inference

$\Rightarrow$ Solving one of the most difficult and fundamental problems in AI and the cognitive sciences

References

Stephen Clarke (2015)
Vector Space Models of Lexical Meaning
Handbook of Contemporary Semantic Theory second edition, edited by Shalom Lappin and Chris Fox Chapter 16, 493 – 522.

Coecke et al. (2010)
Mathematical foundations for a com-positional distributional model of meaning, in J. van Bentham, M. Moortgat, & W. Buszkowski (eds.)
Linguistic Analysis (Lambek Festschrift) volume 36, 345 – 384.

Joachim Lambek (2008)
From Word to Sentence. A Computational Algebraic Ap-proach to Grammar
Polimetrica.

Baroni & Zamparelli (2010)
Nouns are vectors, adjectives are matrices:Representing adjective-noun constructions in semantic space
Conference onEmpirical Methods in Natural Language Processing (EMNLP-10) Cambridge, MA, 1183 – 1193.