

University of Groningen

## Referentiality in individual named event embeddings

Minnema, Gosse; Herbelot, Aurélie

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

### *Document Version*

Final author's version (accepted by publisher, after peer review)

### *Publication date:*

2020

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Minnema, G., & Herbelot, A. (2020). *Referentiality in individual named event embeddings*. Poster session presented at GeCKo Symposium, Barcelona, Spain.

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

### Related work

**Word2Vec**  
"France", "Italy"  
...  
"Paris", "London"

(Gupta et al., EMNLP 2015)

**Freebase**  
GDP  
Population  
Geo-coords  
...

(Bruera 2019, MSc thesis@UniTN)

**Count space**  
"cat", "dog", "carrot"

(Herbelot & Vecchi, EMNLP 2015)

**Quantifiers**  
"X (cat) & is brown(X)"  
"Y (dog) -> is grammatical"  
"X (carrot) & is scaly(X)"

(Kuzmenko & Herbelot, IWCS 2019)

### Theory

**Neo-Davidsonian event semantics**

Napoleon fought yesterday.  
i. fought\_yesterday(napoleon)  
ii.  $\exists x, y. \text{fight}(x) \wedge \text{occurred\_yesterday}(x) \wedge \text{AGENT}(x, \text{napoleon})$

A fight took place yesterday.  
 $\exists x, y. \text{fight}(x) \wedge \text{occurred\_yesterday}(x)$  event sentences  
-> event NPs

**Ontology: events as properties of space-time zones**  
(Bennett 2002)

**FrameNet-inspired semantic roles**  
Fighting, activity, (Combatants, Duration, Manner, Place)

### Battles

**Battle of Waterloo**

### Concert tours

**The Paul McCartney World**

### Conceptual scheme

### Background

- # Formal distributional semantics (combine logical + corpus-based reps)
- # Problem: how to go from vectors to world models?
- # Events are difficult, because (like) individual entities

### Dataset

- # Events with unique names
- # Wiki pages: 1st para + infobox
- # Hurricanes, concerts, battles (n = 1241 / 1978 / 6138)
- # Attributes: year, duration, location, participants ...

### Hurricanes

**Hurricane Isidore**

**UNIVERSITÀ DI TRENTO**

# Gosse Minnema

# Aurélie Herbelot

**university of groningen**

# Referentiality in individual named event embeddings

# How do we talk about events?  
# Compute event representations (BERT, GloVe, Count)  
# Distributional information vs the real world

### Paragraph embeddings

**Idea:** approximate name distribution using content words in definition  
(cf. Lazaridou et al., Cog. Sci. '17; Herbelot & Baroni, EMNLP '17)

**Summed GloVe embeddings**  
word tokens → keep unique content words → retrieve & sum vectors

**BERT embeddings**  
BERT = Google's fancy neural language model  
#1: sentence representation ([CLS] output)  
#2: sum token hidden states (layers 5/9/12)

### Methods

- # Wiki definition embeddings
- # Event name embeddings
- # Attribute prediction
- # Qualitative analysis

### Results

- # Most attributes highly predictable
- # Simple models work well
- # BERT: use individual tokens

### Dimension analysis

Which words are activated by PCA dimensions? (Hurricanes: GloVe)

Dim. 0 (70% of variance) - low activations: "strong", "windings", "faster", "stronger"  
Dim. 1 (24.6% var) - high activations: "places in Indonesia", "intensity", "units"

Dim. 2 (10.17% var) - low / mid activations: "Louisiana", "Texas", "Boston", "County", "Gulf", "Florida", "U.S. geography", "intensity", "units"

Dim. 3 (10.17% var) - low / mid activations: "destroyed", "labor", "3385", "white", "serious", "commensurate", "between", "inductive", "near", "resulted", "step"

**Count dimensions (SVM coefficients)**

Did the battle take place in the eastern (LEFT) or western (RIGHT) hemisphere?

### Event name embeddings

**Count-based**  
Wikipedia corpus → replace event names by token → simple count matrices → PPMI weighting

**'Out-of-the-box'**  
#1: FreeBase W2V (entity name skipgram vecs)  
#2: Wikipedia2Vec (includes graph info)

### Attribute prediction

**Attributes**  
'Semantic roles', based on infobox information (e.g. BATTLE\_YEAR, CONCERT\_TOUR\_DURATION, HURRICANE\_WIND\_SPEED)

Numerical attributes: classes based on frequency distribution (<25th percentile, <50%, <75%, <100%)

**Models**  
SVM (linear) vs. MLP (single hidden layer)  
Separate model for each attribute

### Performance of event description embeddings

### Performance of name count models