Interactive Analysis of Word Vector Embeddings

Florian Heimerl and Michael Gleicher
Department of Computer Sciences
University of Wisconsin - Madison
Interactive Analysis of Word Vector Embeddings

Florian Heimerl and Michael Gleicher
Department of Computer Sciences
University of Wisconsin - Madison
Summary:
Word vector embeddings offer unique challenges

Task analysis of needs

<table>
<thead>
<tr>
<th>characteristic</th>
<th>single target</th>
<th>multiple targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>similarity</td>
<td>(1) inspect local neighborhood</td>
<td>(2) compare local neighborhoods</td>
</tr>
<tr>
<td>average, offset</td>
<td>(3) inspect arithmetic structure</td>
<td>(4) compare arithmetic results</td>
</tr>
<tr>
<td>co-occ. probability</td>
<td>(5) analyze encoded probabilities</td>
<td>(6) compare probabilities</td>
</tr>
<tr>
<td>concept axis</td>
<td>(7) analyze vector relations</td>
<td>(8) compare vector relations</td>
</tr>
<tr>
<td>multiple</td>
<td>(9) discover interesting aspects</td>
<td>(10) compare interesting aspects</td>
</tr>
</tbody>
</table>

3 Designs for unmet needs

Buddy Plots
Co-occurrence Matrices
Concept Axis Plots
What is an embedding?

**General mathematics:**
Place a smaller structure into a larger structure

**Computer science:**
Place a discrete set of objects into a vector space
Encode relationships between objects

High Dimensional Data
Objects have associated Vectors
Kinds of relationships in embeddings

Distance
A is closer to B than to C

Linear Structure
A is to B as C is to D

Semantic Directions
A is more X than C

Relationships are interesting even if global positions are not
Word vector embeddings

Place **words** in a high-dimensional **vector** space

Words **similar in meaning** should be **close in space**

Infer similarity by **distributional semantics**:

- similar context implies similar meaning

Construct embeddings by processing a **corpus** of text

```
my pet cat is brown
my pet dog is brown
my big car is brown
```
Several ways to build embeddings

**Word2Vec**
- Skip-gram model
- Neural embedding

**GLoVE**
- Co-occurrence model
- Factor matrix by optimization
Why use Word Vector Embeddings?

Learn about Language or Corpora (Texts)
Find similar words/synonyms
Track changes of word usage
Exploring polysemy
Creating lexical resources
Evidence of bias
...

Natural Language Applications
Pre-Processing
Translation
Sentiment Analysis
Interpretation
...

...
Challenges of Word Vector Embeddings

Large numbers of Words
High-dimensional spaces
Complex relationships – meaningless positions

Complex processes for building embeddings
Complex downstream applications

No ground truth - subjective aspects
Why visualization?

Tasks are inherently human-centric

Variety of tasks involving interpretation
Linguistic and domain knowledge for applications

But what are those tasks?
Task Analysis: What do people do with Word Embeddings?

Literature survey

111 papers from diverse communities

Consider use cases in Linguistics, HCI, Digital Humanities, etc.

Augment this list by extrapolation

what tasks would users want – but aren’t doing yet
Use cases suggest tasks

Learn about Language or Corpora (Texts)
- Find similar words/synonyms
- Track changes of word usage
- Exploring polysemy
- Creating lexical resources

Natural Language Applications
- Pre-Processing
  - Translation
- Sentiment Analysis
- Interpretation

Interpretation
- Identify items of interest
- Probe values of interest

Evaluation
- Intrinsic (good embedding?)
- Extrinsic (applications success?)
Linguistic Tasks and Characteristics

We identified 7 distinct linguistic tasks within the literature.

Two ways those tasks are relevant:
- Automatically test embeddings through human-curated resources
- Users probe embedding

4 characteristics pertinent to those tasks: similarity, arithmetic structures, concept axis, and co-occurrences

<table>
<thead>
<tr>
<th>linguistic tasks</th>
<th>characteristics</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>rank word pairs</td>
<td>similarity</td>
<td>[BDK14, PSM14]</td>
</tr>
<tr>
<td>compare concepts</td>
<td>average, similarity</td>
<td>[RBS17, SLMJ15]</td>
</tr>
<tr>
<td>find analogies</td>
<td>offset, similarity</td>
<td>[SLMJ15, LG14]</td>
</tr>
<tr>
<td>view neighbors</td>
<td>similarity</td>
<td>[HLJ16, YWL* 16]</td>
</tr>
<tr>
<td>select synonyms</td>
<td>similarity</td>
<td>[BDK14, FDJ* 14]</td>
</tr>
<tr>
<td>project based on concepts</td>
<td>concept axis</td>
<td>[BCZ* 16, FRMW17]</td>
</tr>
<tr>
<td>predict contexts</td>
<td>co-oc. probability</td>
<td>[SN16, LJW* 15]</td>
</tr>
</tbody>
</table>
Prior Work: Word Vector Embedding Visualizations

Liu et al., InfoVis 2017  
Analogy Relationships

Smilkov et al., 2016  
Distance Relationships

Rong and Adar, 2016  
Training Process
Tasks and Characteristics vs. Visualizations

- Rank word pairs
- View neighbors
- Select synonyms
- Compare concepts
- Find analogies
- Project on Concept
- Predict Contexts

- similarity
- similarity
- similarity
- average, similarity
- offset, similarity
- concept axis
- co-oc. probability

Liu, et al 2017
Tasks and Characteristics vs. Visualizations

- Rank word pairs
- View neighbors
- Select synonyms
- Compare concepts
- Find analogies
- Project on Concept
- Predict Contexts

Tasks we seek designs for in this paper

- Similarity
- Average, similarity
- Concept axis
- Co-oc. probability


Liu, et al 2017
1. Similarities (local distances)  

Buddy Plots

2. Co-Occurrances  

Co-occurrence Matrices

3. Concept Axes  

Concept Axis Plots
1. Similarities (local distances)

2. Co-Occurrances

3. Concept Axes
Similarities: Understanding local distances

Distances are meaningful even if absolute values are not

What is close to a word?
Are there groups of words that are similar?

Ordered lists are useful
Density [how many can you show]
Sense of relative distances
Comparison between words
Buddy Plots (1D lists)

Map distance [to selected reference] to horizontal axis

Reference object

Closest to ref

Alexander and Gleicher, 2016 – for Topic Models
Not dimensionality reduction?

Map distance to word to horizontal axis
Focus on a single point – other relations not preserved
Embedding Projector

Nearest points in the original space:
- cattle: 0.543
- cows: 0.556
- dog: 0.607
- dogs: 0.644
- sheep: 0.650
- milk: 0.659
- mad: 0.670
- labyrinth: 0.573
- breed: 0.682
- horse: 0.686
- bull: 0.686
Stacked/Chained buddy plots

Use color to encode distance
Stacked/Chained buddy plots

Use color to encode distance **in the reference row**
Stacked/Chained buddy plots

Use color to encode distance in the reference
Same word... different embedding
1. Similarities (local distances)

2. Co-Occurrances

3. Concept Axes
Why are words similar?
Understanding word co-occurrence

Similarity based on co-occurrence
  count how often one word occurs near another

Co-occurrence matrix
  main form of input data
  many models approximate the matrix [reconstruct]

Useful for understanding and diagnosing models
Co-occurrence view

How do we view the massive matrix?
  - color encoding (heat map) – density, relative values
  - select rows (specify words of interest)
  - select columns (metrics of interestingness – given rows)
    - highest values
    - highest variance
Co-occurrence matrix view

- Co-occurrence matrix for various terms such as 'aunt', 'uncle', 'sister', 'brother', 'woman', and 'man'.
- Each term is represented by a horizontal bar with color gradients indicating frequency or co-occurrence levels.
- Terms like 'son', 'daughter', 'father', 'wife', 'mother', 'married', 'love', 'actress', 'king', 'girl', 'sir', 'female', 'died', 'eldest', 'daughters', 'women', 'nephew', 'duke', 'prince', 'emperor' are listed on the right.
- The color gradient ranges from dark blue (high) to light blue (low).

- The diagram shows the co-occurrence patterns among these terms, with darker colors indicating higher co-occurrence.
Matrix comparison view

High variance words in one embedding may be low in the other
1. Similarities (local distances)

2. Co-Occurrances

3. Concept Axes
Concept Axes: Understanding Semantic Directions

Opposing concepts make an axis
Concept Axes: Understanding Concept Axes

Define an axis from one concept to another
Ways to define axes

**Vector between two concepts**
Interaxis - Kim et al., 2015

**Classifier between two groups**
Explainers – Gleich, 2013
Multiple Concept Axes

Use multi-variate plots

2D = Scatterplot
Scatterplot Interaction
Application: Word meaning change
Application: Stability Assessment
Implementation

http://graphics.cs.wisc.edu/Vis/EmbVis/

Everything runs on line (simplified interfaces)
cloud version uses small models

Python backend / D3 frontend
Limitations

Implementation  Usability and Scalability

Effectiveness  Evaluation [of designs]

Completeness  More Tasks
  Identifying Probes
  Explicit Comparison
  Connection to [model] evaluation
  Feedback to model building
Summary:
Word vector embeddings offer unique challenges

Task analysis of needs

<table>
<thead>
<tr>
<th>characteristic</th>
<th>single target</th>
<th>multiple targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>similarity</td>
<td>(1) inspect local neighborhood</td>
<td>(2) compare local neighborhoods</td>
</tr>
<tr>
<td>average, offset</td>
<td>(3) inspect arithmetic structure</td>
<td>(4) compare arithmetic results</td>
</tr>
<tr>
<td>co-oc. probability</td>
<td>(5) analyze encoded probabilities</td>
<td>(6) compare probabilities</td>
</tr>
<tr>
<td>concept axis</td>
<td>(7) analyze vector relations</td>
<td>(8) compare vector relations</td>
</tr>
<tr>
<td>multiple</td>
<td>(9) discover interesting aspects</td>
<td>(10) compare interesting aspects</td>
</tr>
</tbody>
</table>

Acknowledgments
This work is funded in part by NSF 1162037 and DARPA FA8750-17-2-0107

http://graphics.cs.wisc.edu/Vis/EmbVis/