Alexander Panchenko

INDUCING INTERPRETABLE WORD SENSES FOR WSD AND ENRICHMENT OF LEXICAL RESOURCES
Overview
Inducing word sense representations:

- **word sense embeddings via retrofitting** [Pelevina et al., 2016, Remus & Biemann, 2018];
- **inducing synsets** [Ustalov et al., 2017b, Ustalov et al., 2017a, Ustalov et al., 2018]
- **inducing semantic classes** [Panchenko et al., 2018]
Inducing word sense representations:

- word sense embeddings via retrofitting [Pelevina et al., 2016, Remus & Biemann, 2018];
- inducing synsets [Ustalov et al., 2017b, Ustalov et al., 2017a, Ustalov et al., 2018]
- inducing semantic classes [Panchenko et al., 2018]

Making induced senses interpretable [Panchenko et al., 2017b, Panchenko et al., 2017c]
Inducing word sense representations:
- word sense embeddings via retrofitting [Pelevina et al., 2016, Remus & Biemann, 2018];
- inducing synsets [Ustalov et al., 2017b, Ustalov et al., 2017a, Ustalov et al., 2018]
- inducing semantic classes [Panchenko et al., 2018]

Making induced senses interpretable [Panchenko et al., 2017b, Panchenko et al., 2017c]

Linking induced word senses to lexical resources [Panchenko, 2016, Faralli et al., 2016, Panchenko et al., 2017a, Biemann et al., 2018]
Inducing word sense representations
Inducing word sense representations

Word vs sense embeddings

- armchair
- chair
- furniture
- plate
- tray
- table
- plot
- graph
- column
- data
- excel
Inducing word sense representations

Word vs sense embeddings
Related work

- **Sense Embeddings**
  - **Knowledge-Based:** use inventory of a lexical resource
    - Chen et al. (2014)
    - Rothe & Schütze (2015)
    - Camacho-Collados et al. (2015)
    - Iacobacci et al. (2015)
    - Nieto Pina and Johansson (2016)
  - **Knowledge-Free:** Induce an inventory from text corpus
    - Induce sense embeddings directly
      - Huang et al. (2012)
      - Tian et al. (2014)
      - Neelakantan et al. (2014)
      - Li and Jurafsky (2015)
      - Bartunov et al. (2016)
      - Lee and Chen (2017)
    - Retrofit word embeddings into an induced inventory
      - Pelevina et al. (2016)
      - Remus and Biemann (2018)
**AutoExtend** [Rothe & Schütze, 2015]

* image is reproduced from the original paper
Adagram [Bartunov et al., 2016]

Multiple vector representations $\theta$ for each word:

$\text{Adagram}$ [Bartunov et al., 2016]

Multiple vector representations $\theta$ for each word:
Related work: knowledge-free

- **Adagram** [Bartunov et al., 2016]
- Multiple vector representations $\theta$ for each word:

$$ p(Y, Z, \beta|X, \alpha, \theta) = \prod_{w=1}^{V} \prod_{k=1}^{\infty} p(\beta_{wk}|\alpha) \prod_{i=1}^{N} [p(z_i|x_i, \beta) \prod_{j=1}^{C} p(y_{ij}|z_i, x_i, \theta)], $$

- $z_i$ – a hidden variable: a sense index of word $x_i$ in context $C$;
- $\alpha$ – a meta-parameter controlling number of senses.
Related work: knowledge-free

- **Adagram** [Bartunov et al., 2016]
- Multiple vector representations $\theta$ for each word:

$$p(Y, Z, \beta | X, \alpha, \theta) = \prod_{w=1}^{V} \prod_{k=1}^{\infty} p(\beta_{wk} | \alpha) \prod_{i=1}^{N} [p(z_i | x_i, \beta) \prod_{j=1}^{C} p(y_{ij} | z_i, x_i, \theta)],$$

- $z_i$ – a hidden variable: a sense index of word $x_i$ in context $C$;
- $\alpha$ – a meta-parameter controlling number of senses.

- **See also**: [Neelakantan et al., 2014] and [Li and Jurafsky, 2015]
Word sense induction (WSI) based on **graph clustering**:
- [Lin, 1998]
- [Pantel and Lin, 2002]
- [Widdows and Dorow, 2002]
- **Chinese Whispers** [Biemann, 2006]
- [Hope and Keller, 2013]
Inducing word sense representations

Related work: Chinese Whispers#1

* source of the image: http://ic.pics.livejournal.com/blagin_anton/33716210/2701748/2701748_800.jpg
Related work: Chinese Whispers#2
Inducing word sense representations

Related work: Chinese Whispers#2
Related work: Chinese Whispers#2
Inducing word sense representations

**Sense embeddings using retrofitting**

RepL4NLP@ACL’16 [Pelevina et al., 2016], LREC’18 [Remus & Biemann, 2018]

**Prior methods:**
- Induce inventory by **clustering of word instances**
- Use **existing** sense inventories

**Our method:**
- **Input:** word embeddings
- **Output:** word sense embeddings
- **Word sense induction** by **clustering of word ego-networks**
From word embeddings to sense embeddings

1. Learning Word Vectors
2. Calculate Word Similarity Graph
3. Word Sense Induction
4. Pooling of Word Vectors

Text Corpus → Sense Vectors

Calculation of Word Similarity Graph → Word Vectors

Word Vectors → Word Similarity Graph

Sense Inventory → Pooling of Word Vectors
- Word sense induction using ego-network clustering
Neighbours of Word and Sense Vectors

<table>
<thead>
<tr>
<th>Vector</th>
<th>Nearest Neighbors</th>
</tr>
</thead>
<tbody>
<tr>
<td>table</td>
<td>tray, bottom, diagram, bucket, brackets, stack, basket, list, parenthesis, cup, saucer, pile, playfield, bracket, pot, drop-down, cue, plate</td>
</tr>
</tbody>
</table>
## Neighbours of Word and Sense Vectors

<table>
<thead>
<tr>
<th>Vector</th>
<th>Nearest Neighbors</th>
</tr>
</thead>
<tbody>
<tr>
<td>table</td>
<td>tray, bottom, diagram, bucket, brackets, stack, basket, list, parenthesis, cup, saucer, pile, playfield, bracket, pot, drop-down, cue, plate</td>
</tr>
<tr>
<td>table#0</td>
<td>leftmost#0, column#1, tableau#1, indent#1, bracket#3, pointer#0, footer#1, cursor#1, diagram#0, grid#0</td>
</tr>
<tr>
<td>table#1</td>
<td>pile#1, stool#1, tray#0, basket#0, bowl#1, bucket#0, box#0, cage#0, saucer#3, mirror#1, pan#1, lid#0</td>
</tr>
</tbody>
</table>
Word Sense Disambiguation

1. **Context extraction**: use context words around the target word

2. **Context filtering**: based on context word’s relevance for disambiguation

3. **Sense choice in context**: maximise similarity between a context vector and a sense vector
Inducing word sense representations

Sense embeddings using retrofitting

They bought a **table** and chairs for kitchen.

<table>
<thead>
<tr>
<th>table senses</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
</tr>
<tr>
<td>furniture</td>
</tr>
</tbody>
</table>
Inducing word sense representations using retrofitting.

Sense embeddings using retrofitting.

They bought a **table** and chairs for kitchen.

Table senses: data, furniture.

Context window = 4.

They bought a **table** and chairs for kitchen.
Inducing word sense representations

Sense embeddings using retrofitting

They bought a table and chairs for kitchen.

context window = 4

context filtering

chairs + kitchen

table senses

data furniture

They bought a table and chairs for kitchen.
Inducing word sense representations

Sense embeddings using retrofitting

They bought a table and chairs for kitchen.

Context window = 4

Cosine ($c_j, s_i$)

Relevance score

Context filtering

Cosine ($c_{avg}, s_i$)

Table (furniture)
Unsupervised WSD SemEval’13, ReprL4NLP [Pelevina et al., 2016]:
- comparable to SOTA, incl. Adagram sense embeddings.
Inducing word sense representations

Sense embeddings using retrofitting

**Unsupervised WSD** SemEval’13, ReprL4NLP [Pelevina et al., 2016]:
- comparable to SOTA, incl. Adagram sense embeddings.

**Semantic relatedness**, LREC’2018 [Remus & Biemann, 2018]:

<table>
<thead>
<tr>
<th></th>
<th>AUTOEXTEND</th>
<th>ADAGRAM</th>
<th>SGNS</th>
<th>GLOVE</th>
<th>SYMPAT</th>
<th>LSABOW</th>
<th>LSAHAL</th>
<th>PARAGRAMSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMLEX999</td>
<td>0.45</td>
<td>0.29</td>
<td>0.44</td>
<td>0.37</td>
<td>0.54</td>
<td>0.30</td>
<td>0.27</td>
<td>0.68</td>
</tr>
<tr>
<td>MEN</td>
<td>0.72</td>
<td>0.67</td>
<td>0.77</td>
<td>0.73</td>
<td>0.53</td>
<td>0.67</td>
<td>0.71</td>
<td>0.77</td>
</tr>
<tr>
<td>SIMVERB</td>
<td>0.43</td>
<td>0.27</td>
<td>0.36</td>
<td>0.23</td>
<td>0.37</td>
<td>0.15</td>
<td>0.19</td>
<td>0.53</td>
</tr>
<tr>
<td>WORDSIM353</td>
<td>0.58</td>
<td>0.61</td>
<td>0.70</td>
<td>0.61</td>
<td>0.47</td>
<td>0.67</td>
<td>0.59</td>
<td>0.72</td>
</tr>
<tr>
<td>SIMLEX999-N</td>
<td>0.44</td>
<td>0.33</td>
<td>0.45</td>
<td>0.39</td>
<td>0.48</td>
<td>0.32</td>
<td>0.34</td>
<td>0.68</td>
</tr>
<tr>
<td>MEN-N</td>
<td>0.72</td>
<td>0.68</td>
<td>0.77</td>
<td>0.76</td>
<td>0.57</td>
<td>0.71</td>
<td>0.73</td>
<td>0.78</td>
</tr>
</tbody>
</table>
Unsupervised WSD SemEval’13, ReprL4NLP [Pelevina et al., 2016]:
- comparable to SOTA, incl. sense embeddings.

Semantic relatedness, LREC’2018 [Remus & Biemann, 2018]:

<table>
<thead>
<tr>
<th>Method</th>
<th>AUTOEXTEND</th>
<th>ADAGRAM</th>
<th>SGNS</th>
<th>SGNS+Senses</th>
<th>GLOVE</th>
<th>GLOVE+Senses</th>
<th>SYMPAT</th>
<th>SYMPAT+Senses</th>
<th>LSABOW</th>
<th>LSABOW+Senses</th>
<th>LSAHAL</th>
<th>LSAHAL+Senses</th>
<th>PARAGRAMSL</th>
<th>PARAGRAMSL+Senses</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMLEX999</td>
<td>0.45</td>
<td>0.29</td>
<td>0.44</td>
<td>0.46</td>
<td>0.37</td>
<td>0.41</td>
<td>0.54</td>
<td>0.55</td>
<td>0.30</td>
<td>0.39</td>
<td>0.27</td>
<td>0.38</td>
<td>0.68</td>
<td>0.64</td>
</tr>
<tr>
<td>MEN</td>
<td>0.72</td>
<td>0.67</td>
<td>0.77</td>
<td>0.78</td>
<td>0.73</td>
<td>0.77</td>
<td>0.53</td>
<td>0.68</td>
<td>0.67</td>
<td>0.70</td>
<td>0.71</td>
<td>0.74</td>
<td>0.77</td>
<td>0.80</td>
</tr>
<tr>
<td>SIMVERB</td>
<td>0.43</td>
<td>0.27</td>
<td>0.36</td>
<td>0.39</td>
<td>0.23</td>
<td>0.30</td>
<td>0.37</td>
<td>0.45</td>
<td>0.15</td>
<td>0.22</td>
<td>0.19</td>
<td>0.28</td>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>WORDSIM353</td>
<td>0.58</td>
<td>0.61</td>
<td>0.70</td>
<td>0.69</td>
<td>0.61</td>
<td>0.65</td>
<td>0.47</td>
<td>0.62</td>
<td>0.67</td>
<td>0.66</td>
<td>0.59</td>
<td>0.63</td>
<td>0.72</td>
<td>0.73</td>
</tr>
<tr>
<td>SIMLEX999-N</td>
<td>0.44</td>
<td>0.33</td>
<td>0.45</td>
<td>0.50</td>
<td>0.39</td>
<td>0.47</td>
<td>0.48</td>
<td>0.55</td>
<td>0.32</td>
<td>0.46</td>
<td>0.34</td>
<td>0.44</td>
<td>0.68</td>
<td>0.66</td>
</tr>
<tr>
<td>MEN-N</td>
<td>0.72</td>
<td>0.68</td>
<td>0.77</td>
<td>0.79</td>
<td>0.76</td>
<td>0.80</td>
<td>0.57</td>
<td>0.74</td>
<td>0.71</td>
<td>0.73</td>
<td>0.73</td>
<td>0.76</td>
<td>0.78</td>
<td>0.81</td>
</tr>
</tbody>
</table>
Inducing word sense representations using retrofitting.

**Sense embeddings using retrofitting**

- Word and sense embeddings of words **iron** and **vitamin**.

**LREC’18** [Remus & Biemann, 2018]
### ACL’17 [Ustalov et al., 2017b]

Examples of extracted synsets:

<table>
<thead>
<tr>
<th>Size</th>
<th>Synset</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>{decimal point, dot}</td>
</tr>
<tr>
<td>3</td>
<td>{gullet, throat, food pipe}</td>
</tr>
<tr>
<td>4</td>
<td>{microwave meal, ready meal, TV dinner, frozen dinner}</td>
</tr>
<tr>
<td>5</td>
<td>{objective case, accusative case, oblique case, object case, accusative}</td>
</tr>
<tr>
<td>6</td>
<td>{radio theater, dramatized audiobook, audio theater, radio play, radio drama, audio play}</td>
</tr>
</tbody>
</table>
Outline of the ’Watset’ method:
Inducing word sense representations

Synset induction

Stage 1: Ambiguous Graph before the Local Clustering
Stage 2: Sense Inventory with Ambiguous Neighbors

- bank^2
- building?
- bank building?
- streambank?
- streamside^1
- riverbank?
Stage 3: Disambiguated Graph before the Global Clustering
Inducing interpretable word senses for WSD and enrichment of lexical resources

Synset induction
Inducing Interpretable Word Senses for WSD and Enrichment of Lexical Resources, Alexander Panchenko

Synset induction

F-score comparison across different word sense induction methods for WordNet (English), BabelNet (English), RuWordNet (Russian), and YARN (Russian).
### Examples of semantic classes:

<table>
<thead>
<tr>
<th>ID</th>
<th>Sense Cluster</th>
<th>Hypernyms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>peach#1, banana#1, pineapple#0, berry#0, blackberry#0, grapefruit#0, strawberry#0, blueberry#0, grape#0, melon#0, orange#0, pear#0, plum#0, raspberry#0, watermelon#0, apple#0, apricot#0, ...</td>
<td>fruit#0, crop#0, ingredient#0, food#0, ...</td>
</tr>
<tr>
<td>2</td>
<td>C#4, Basic#2, Haskell#5, Flash#1, Java#1, Pascal#0, Ruby#6, PHP#0, Ada#1, Oracle#3, Python#3, Apache#3, Visual Basic#1, ASP#2, Delphi#2, SQL Server#0, CSS#0, AJAX#0, the Java#0, ...</td>
<td>programming language#3, technology#0, language#0, format#2, app#0</td>
</tr>
</tbody>
</table>
Induction of semantic classes

Inducing word sense representations

Noisy Hypernyms → Cleansed Hypernyms

Text Corpus

Induction of Semantic Classes

Word Sense Induction from Text Corpus → Representing Senses with Ego Networks → Sense Graph Construction → Clustering of Word Senses

Induced Word Senses → Sense Ego-Networks → Global Sense Graph

Labeling Sense Clusters with Hypernyms → Noisy Hypernyms → Cleansed Hypernyms

Global Sense Clusters

Semantic Classes
Filtering noisy hypernyms with semantic classes

LREC’18 [Panchenko et al., 2018]:

\[
\text{Hypernyms, } \mathcal{H}(c) \subset S
\]

\[
\begin{align*}
\text{city} & \#2 \\
\text{food} & \#0 \\
\text{fruit} & \#1 \\
\text{apple} & \#2 \\
\text{mango} & \#0 \\
\text{pear} & \#0 \\
\text{mangosteen} & \#0
\end{align*}
\]

\[
\text{Sense Cluster, } c \subset S
\]
Filtering of a noisy hypernymy database with semantic classes. 

**LREC’18** [Panchenko et al., 2018]

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Hypernyms</td>
<td>0.475</td>
<td>0.546</td>
<td>0.508</td>
</tr>
<tr>
<td>(Seitner et al., 2016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic Classes</td>
<td>0.541</td>
<td>0.679</td>
<td>0.602</td>
</tr>
<tr>
<td>(coarse-grained)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Making induced senses interpretable
Knowledge-based sense representations are interpretable.

Python (programming language) /
/usr/bin/python
/usr/local/bin/python Python language Python programming language

Python is a widely used general-purpose, high-level programming language. Wikipedia

More definitions

More relations

EXPLORE NETWORK
Most **knowledge-free** sense representations are **uninterpretable**

```
In [11]: sv.syn0[sv.vocab["python#2"].index]
Out[11]:
array([-0.0493343 , -0.02244579,  0.02296794,  0.03484775,  0.0404554 ,
       0.04304857, -0.02211852, -0.02118347, -0.03212074, -0.01202453,
       0.01206081,  0.05609602, -0.05950832,  0.00859888, -0.01051112,
       0.03177784, -0.06489294,  0.03833736,  0.05437034, -0.01451268,
       -0.02419239, -0.03195219,  0.0620546 ,  0.10284331,  0.07430374,
       -0.04109243, -0.01181133 ,  0.05401124,  0.05283536,  0.00873093,
       0.03662092,  0.03762468,  0.02368712, -0.03980339,  0.02791001,
       0.02529952, -0.02255581, -0.00925604, -0.03940469, -0.02855149,
         ...          -0.08179335,  0.02319797, -0.0167018 ,  0.04818865, -0.06946786,
       -0.06530198,  0.00522405, -0.0336296 , -0.05401101,  0.01190361], dtype=float32)
```
Making induced senses interpretable

http://jobimtext.org/wsd

Predicted senses for 'Jaguar'

1. jaguar (animal)
   Similarity score: 0.00184 / Confidence: 99.87% / Sense ID: jaguar#0 / BabelNet ID: bn:00033987n

   Hypnyms: animal, wildlife, bird, mammal

   Sample sentences:
   The jaguar, a compact and well-muscled animal, is the largest cat in the New World.
   Jaguar may leap onto the back of the prey and sever the cervical vertebrae, immobilizing the target.

   Cluster words: lion, tiger, leopard, wolf, monkey, otter, crocodile, alligator, deer, cat, elephant, fox, eagle, owl, snake

   Context words:
   elephant: 0.012, tiger: 0.012, fox: 0.0099, wolf: 0.0097, cub: 0.0086, monkey: 0.0083, leopard: 0.0074, eagle: 0.0062
   den: 0.0043, elk: 0.0040, 32078 more not shown

   Matching features:
   leopard: 0.0011, predator: 0.00040, spotted: 0.00038, large: 0.0000041, similar: 0.0000015, tropical: 5.6e-7, america: 2.0e-7
Making induced senses interpretable

Making induced senses interpretable

Hypernymy prediction in context. **EMNLP’17** [Panchenko et al., 2017b]
11,702 sentences, 863 words with **avg. polysemy of 3.1**.

<table>
<thead>
<tr>
<th>WSD Model</th>
<th>Inventory</th>
<th>Features</th>
<th>Accuracy Hypers</th>
<th>Accuracy HyperHypers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Senses</td>
<td>Random</td>
<td></td>
<td>0.257</td>
<td>0.610</td>
</tr>
<tr>
<td>Word Senses</td>
<td>MFS</td>
<td></td>
<td>0.292</td>
<td>0.682</td>
</tr>
<tr>
<td>Word Senses</td>
<td>Cluster Words</td>
<td></td>
<td>0.291</td>
<td>0.650</td>
</tr>
<tr>
<td>Word Senses</td>
<td>Context Words</td>
<td></td>
<td><strong>0.308</strong></td>
<td><strong>0.686</strong></td>
</tr>
</tbody>
</table>
11,702 sentences, 863 words with **avg.polysemy of 3.1**.

<table>
<thead>
<tr>
<th><strong>WSD Model</strong></th>
<th><strong>Inventory</strong></th>
<th><strong>Features</strong></th>
<th><strong>Accuracy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word Senses</strong></td>
<td>Random</td>
<td>0.257</td>
<td>0.610</td>
</tr>
<tr>
<td><strong>Word Senses</strong></td>
<td>MFS</td>
<td>0.292</td>
<td>0.682</td>
</tr>
<tr>
<td><strong>Word Senses</strong></td>
<td>Cluster Words</td>
<td>0.291</td>
<td>0.650</td>
</tr>
<tr>
<td><strong>Word Senses</strong></td>
<td>Context Words</td>
<td><strong>0.308</strong></td>
<td><strong>0.686</strong></td>
</tr>
<tr>
<td><strong>Super Senses</strong></td>
<td>Random</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Super Senses</strong></td>
<td>MFS</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Super Senses</strong></td>
<td>Cluster Words</td>
<td><strong>0.174</strong></td>
<td><strong>0.365</strong></td>
</tr>
<tr>
<td><strong>Super Senses</strong></td>
<td>Context Words</td>
<td>0.086</td>
<td>0.188</td>
</tr>
</tbody>
</table>
Linking induced senses to resources
Inducing Interpretable Word Senses for WSD and Enrichment of Lexical Resources

Construction of Proto-Conceptualization (PCZ)

- Induce a Graph of Sem. Related Words
- Word Sense Induction
- Labeling Senses with Hypernyms
- Disambiguation of Neighbours
- Construction of sense feature representations

Linking Proto-Conceptualization to Lexical Resource

- Typing of the Unmapped Induced Senses
- Linking Induced Senses to Senses of the LR
- Part. Linked Senses to the LR

Text Corpus

Enriched Lexical Resource

Lexical Resource (LR): WordNet, BabelNet, ...

LREC’16 [Panchenko, 2016], ISWC’16 [Faralli et al., 2016], SENSE@EACL’17 [Panchenko et al., 2017a], NLE’18 [Biemann et al., 2018]
### Linking induced senses to resources

<table>
<thead>
<tr>
<th>Word</th>
<th>AdaGram</th>
<th>BabelNet</th>
<th>AdaGram BoW</th>
<th>BabelNet BoW</th>
</tr>
</thead>
<tbody>
<tr>
<td>python</td>
<td>2</td>
<td>bn:01713224n</td>
<td>perl, php, java, smalltalk, ruby, lua, tcl, scripting, javascript, bindings, binding, programming, coldfusion, actionscript, net, . . .</td>
<td>language, programming, pythonista, python programming, python3, python2, level, computer, pythonistas, python3000, . . .</td>
</tr>
<tr>
<td>python</td>
<td>1</td>
<td>bn:01157670n</td>
<td>monty, circus, spamalot, python, magoo, muppet, snoopy, featurette, disney, tunes, tune, classic, shorts, short, apocalypse, . . .</td>
<td>monty, comedy, monty python, british, monte, monte python, troupe, pythonesque, foot, artist, record, surreal, terry, . . .</td>
</tr>
<tr>
<td>python</td>
<td>3</td>
<td>bn:00046456n</td>
<td>spectacled, unicorns, snake, giant, caiman, leopard, squirrel, crocodile, horned, cat, mole, elephant, opossum, pheasant, . . .</td>
<td>molurus, indian, boa, tigris, tiger python, rock, tiger, indian python, reptile, python molurus, indian rock python, coluber, . . .</td>
</tr>
<tr>
<td>python</td>
<td>4</td>
<td>bn:01157670n</td>
<td>circus, fly, flying, dusk, lizard, moth, unicorn, puff, adder, vulture, tyrannosaurus, zephyr, badger, . . .</td>
<td>monty, comedy, monty python, british, monte, monte python, troupe, pythonesque, foot, artist, record, surreal, terry, . . .</td>
</tr>
<tr>
<td>python</td>
<td>1</td>
<td>bn:00473212n</td>
<td>monty, circus, spamalot, python, magoo, muppet, snoopy, featurette, disney, tunes, tune, classic, shorts, short, apocalypse, . . .</td>
<td>pictures, monty, python monty pictures, limited, company, python pictures limited, kingdom, picture, serve, director, . . .</td>
</tr>
<tr>
<td>python</td>
<td>1</td>
<td>bn:03489893n</td>
<td>monty, circus, spamalot, python, magoo, muppet, snoopy, featurette, disney, tunes, tune, classic, shorts, short, apocalypse, . . .</td>
<td>film, horror, movie, clabaugh, richard, monster, century, direct, snake, python movie, television, giant, natural, language, for-tv, . . .</td>
</tr>
<tr>
<td>Model</td>
<td>Representation of the Sense &quot;disk (medium)&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WordNet</td>
<td>memory, device, floppy, disk, hard, disk, disk, computer, science, computing, diskette, fixed, disk, floppy, magnetic, disc, magnetic, disk, hard, disc, storage, device</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WordNet + Linked</td>
<td>recorder, disk, floppy, console, diskette, handset, desktop, iphone, iPod, HDTV, kit, RAM, Discs, Blu-ray, computer, GB, microchip, site, cartridge, printer, tv, VCR, Disc, player, LCD, software, component, camcorder, cellphone, card, monitor, display, burner, Web, stereo, internet, model, itunes, turntable, chip, cable, camera, iphone, notebook, device, server, surface, wafer, page, drive, laptop, screen, pc, television, hardware, YouTube, dvr, DVD, product, folder, VCR, radio, phone, circuitry, partition, megabyte, peripheral, format, machine, tuner, website, merchandise, equipment, gb, discs, MP3, hard-drive, piece, video, storage device, memory device, microphone, hd, EP, content, soundtrack, webcam, system, blade, graphic, microprocessor, collection, document, programming, battery, keyboard, HD, handheld, CDs, reel, web, material, hard-disk, ep, chart, debut, configuration, recording, album, broadcast, download, fixed disk, planet, pda, microfilm, iPod, videotape, text, cylinder, cpu, canvas, label, sampler, workstation, electrode, magnetic disc, catheter, magnetic disk, Video, mobile, cd, song, modern, mouse, tube, set, ipad, signal, substrate, vinyl, music, clip, pad, audio, compilation, memory, message, reissue, ram, CD, subsystem, hdd, touchscreen, electronics, demo, shell, sensor, file, shelf, processor, cassette, extra, mainframe, motherboard, floppy disk, lp, tape, version, kilobyte, pacemaker, browser, Playstation, pager, module, cache, DVD, movie, Windows, cd-rom, e-book, valve, directory, harddrive, smartphone, audiotape, technology, hard disk, show, computing, computer science, Blu-Ray, blu-ray, HDD, HD-DVD, scanner, hard disc, gadget, booklet, copier, playback, TiVo, controller, filter, DVDs, gigabyte, paper, mp3, CPU, dvd-r, pipe, cd-r, playlist, slot, VHS, film, videocassette, interface, adapter, database, manual, book, channel, changer, storage</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evaluation of linking accuracy:

(a) Precision vs. Recall

- global threshold, cosine (AUC = 0.78)
- random (AUC = 0.11)

(b) Recall vs. Threshold

- red: recall
- blue: precision
Evaluation of enriched representations based on WSD:

- SensEval-3 coarse-grained
  - WordNet
  - WordNet + Related (news)
  - WordNet + Related (news) + Context (wiki)
  - WordNet + Related (news) + Context (news)

- SemEval-2007 fine-grained
  - WordNet
  - WordNet + Related (news)
  - WordNet + Related (news) + Context (wiki)
  - WordNet + Related (news) + Context (news)
Conclusion
Conclusion

Vectors + Graphs = ♡

GRAPHS

NOT DEAD.
We can **induce word senses, synsets** and **semantic classes** in a knowledge-free way using **graph clustering** and **distributional models**.
We can **induce word senses, synsets and semantic classes** in a knowledge-free way using **graph clustering** and **distributional models**.

We can make the **induced word senses interpretable** in a knowledge-free way with **hypernyms, images, definitions**.
We can **induce word senses, synsets** and **semantic classes** in a knowledge-free way using **graph clustering** and distributional models.

We can make the **induced word senses interpretable** in a knowledge-free way with **hypernyms, images, definitions**.

We can **link induced senses to lexical resources** to
- improve **performance of WSD**;
- **enrich lexical resources** with emerging senses.
An ongoing shared task on WSI&D

- Participate in an **ACL SIGSLAV** sponsored shared task on **word sense induction and disambiguation** for Russian!

Thank you! Questions?

This research was supported by

DFG Deutsche Forschungsgemeinschaft

Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

DAAD
<table>
<thead>
<tr>
<th>Model</th>
<th>Jacc.</th>
<th>Tau</th>
<th>WNDCG</th>
<th>F.NMI</th>
<th>F.B-Cubed</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI-KU (add1000)</td>
<td>0.176</td>
<td>0.609</td>
<td>0.205</td>
<td>0.033</td>
<td>0.317</td>
</tr>
<tr>
<td>AI-KU</td>
<td>0.176</td>
<td>0.619</td>
<td>0.393</td>
<td>0.066</td>
<td>0.382</td>
</tr>
<tr>
<td>AI-KU (remove5-add1000)</td>
<td>0.228</td>
<td>0.654</td>
<td>0.330</td>
<td>0.040</td>
<td>0.463</td>
</tr>
<tr>
<td>Unimelb (5p)</td>
<td>0.198</td>
<td>0.623</td>
<td>0.374</td>
<td>0.056</td>
<td>0.475</td>
</tr>
<tr>
<td>Unimelb (50k)</td>
<td>0.198</td>
<td>0.633</td>
<td>0.384</td>
<td>0.060</td>
<td>0.494</td>
</tr>
<tr>
<td>UoS (#WN senses)</td>
<td>0.171</td>
<td>0.600</td>
<td>0.298</td>
<td>0.046</td>
<td>0.186</td>
</tr>
<tr>
<td>UoS (top-3)</td>
<td>0.220</td>
<td>0.637</td>
<td>0.370</td>
<td>0.044</td>
<td>0.451</td>
</tr>
<tr>
<td>La Sapienza (1)</td>
<td>0.131</td>
<td>0.544</td>
<td>0.332</td>
<td>0.1 -</td>
<td>0.1 -</td>
</tr>
<tr>
<td>La Sapienza (2)</td>
<td>0.131</td>
<td>0.535</td>
<td>0.394</td>
<td>0.1 -</td>
<td>0.1 -</td>
</tr>
<tr>
<td>AdaGram, $\alpha = 0.05$, 100 dim</td>
<td>0.274</td>
<td>0.644</td>
<td>0.318</td>
<td>0.058</td>
<td>0.470</td>
</tr>
<tr>
<td>w2v</td>
<td>0.197</td>
<td>0.615</td>
<td>0.291</td>
<td>0.011</td>
<td>0.615</td>
</tr>
<tr>
<td>w2v (nouns)</td>
<td>0.179</td>
<td>0.626</td>
<td>0.304</td>
<td>0.011</td>
<td>0.623</td>
</tr>
<tr>
<td>JBT</td>
<td>0.205</td>
<td>0.624</td>
<td>0.291</td>
<td>0.017</td>
<td>0.598</td>
</tr>
<tr>
<td>JBT (nouns)</td>
<td>0.198</td>
<td>0.643</td>
<td>0.310</td>
<td>0.031</td>
<td>0.595</td>
</tr>
<tr>
<td>TWSI (nouns)</td>
<td>0.215</td>
<td>0.651</td>
<td>0.318</td>
<td>0.030</td>
<td>0.573</td>
</tr>
</tbody>
</table>


Unsupervised does not mean uninterpretable: The case for word sense induction and disambiguation.


Retrofitting word representations for unsupervised sense aware word similarities.


In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 94–105).: Springer.
Watset: Automatic induction of synsets from a graph of synonyms.

Word sense disambiguation based on automatically induced synsets.
Accepted for publication.