Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Alexander Panchenko

INDUCING INTERPRETABLE WORD SENSES FOR WSD AND ENRICHMENT OF LEXICAL RESOURCES

Overview

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]

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

Overview

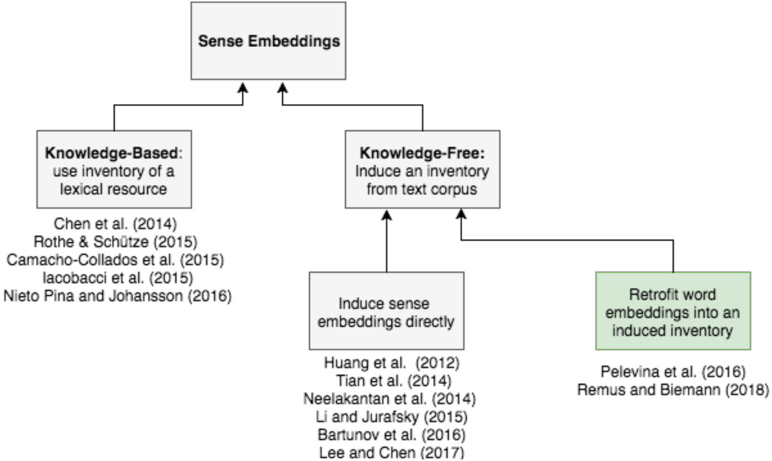
- **Inducing word sense representations**:
  - **word sense embeddings via retrofitting** [Pelevina et al., 2016, Remus & Biemann, 2018];
  - **inducing synsets** [Ustalov et al., 2017b, Ustalov et al., 2017a, Ustalov et al., 2018]
  - **inducing semantic classes** [Panchenko et al., 2018]
- **Making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

Word vs sense embeddings

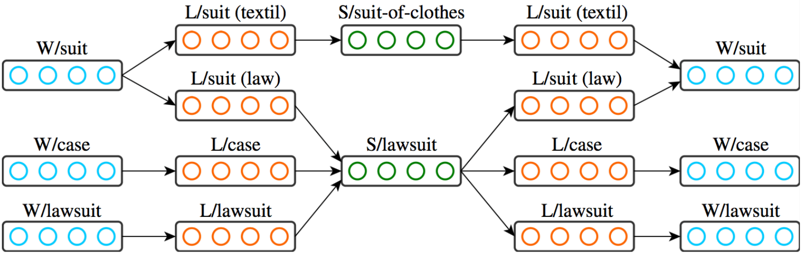
Word vs sense embeddings

Related work



Related work: knowledge-based

- **AutoExtend** [Rothe & Schütze, 2015]



* image is reproduced from the original paper

Related work: knowledge-free

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

Related work: knowledge-free

- **Adagram** [Bartunov et al., 2016]
- Multiple vector representations θ 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;
  - α – a meta-parameter controlling number of senses.

Related work: knowledge-free

- **Adagram** [Bartunov et al., 2016]
- Multiple vector representations θ 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;
  - α – a meta-parameter controlling number of senses.
- **See also**: [Neelakantan et al., 2014] and [Li and Jurafsky, 2015]

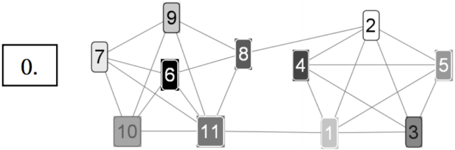
Related work: word sense induction

- 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]

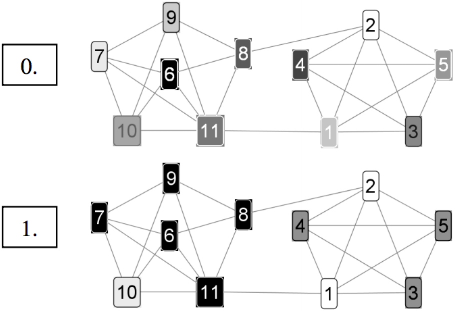
Related work: Chinese Whispers#1

* source of the image: `http://ic.pics.livejournal.com/blagin_anton/33716210/2701748/2701748_800.jpg`

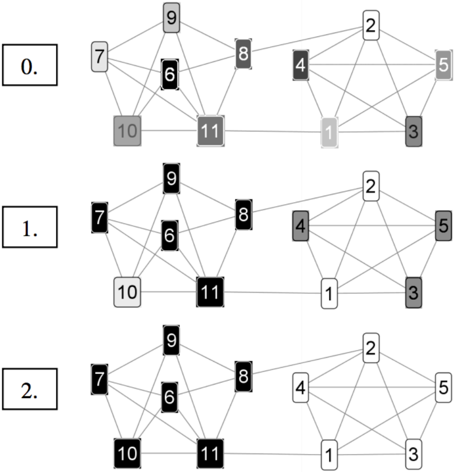
Related work: Chinese Whispers#2



Related work: Chinese Whispers#2



Related work: Chinese Whispers#2



Sense embeddings using retrofitting

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

Prior methods:

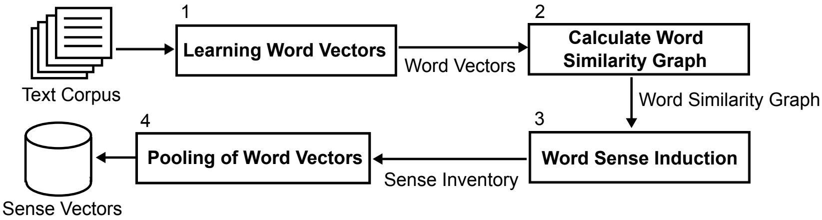
- Induce inventory by clustering of word instances
- Use existing sense inventories

Our method:

- **Input:** word embeddings
- **Output:** word sense embeddings
- **Word sense induction** by clustering of word ego-networks

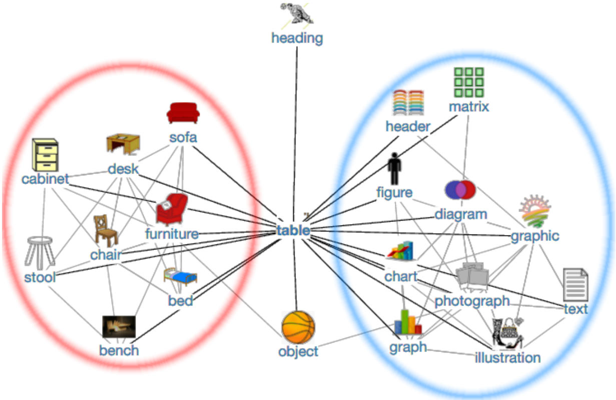
Sense embeddings using retrofitting

- From word embeddings to sense embeddings



Sense embeddings using retrofitting

- Word sense induction using ego-network clustering



Sense embeddings using retrofitting

- Neighbours of Word and Sense Vectors

Vector	Nearest Neighbors
table	tray, bottom, diagram, bucket, brackets, stack, basket, list, parenthesis, cup, saucer, pile, playfield, bracket, pot, drop-down, cue, plate

Sense embeddings using retrofitting

- Neighbours of Word and Sense Vectors

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

Sense embeddings using retrofitting

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

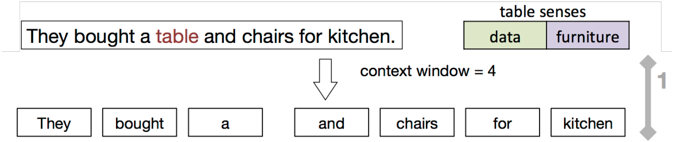
Sense embeddings using retrofitting

They bought a table and chairs for kitchen.

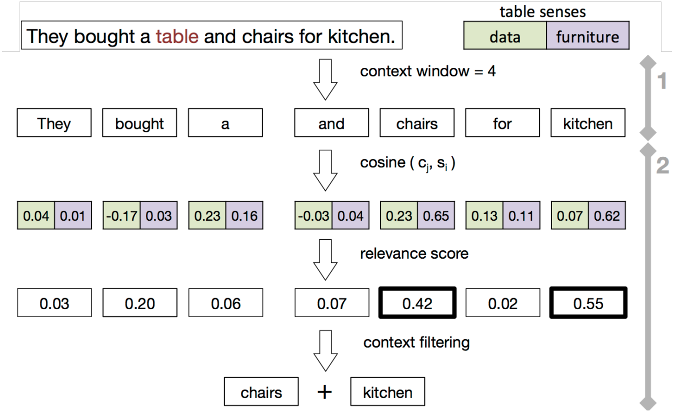
table senses

data	furniture

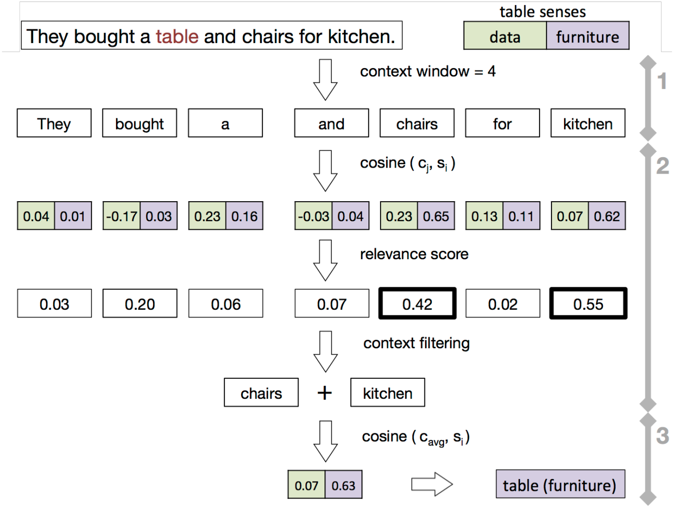
Sense embeddings using retrofitting



Sense embeddings using retrofitting



Sense embeddings using retrofitting



Sense embeddings using retrofitting

Unsupervised WSD SemEval'13, **ReprL4NLP** [Pelevina et al., 2016]:

- comparable to SOTA, incl. sense embeddings.

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

	AUTOEXTEND	ADAGRAM	SGNS	.	GLOVE	.	SYMPAT	.	LSABOW	.	LSAHAL	.	PARAGRAMSL	.
SIMLEX999	0.45	0.29	0.44		0.37		0.54		0.30		0.27		0.68	
MEN	0.72	0.67	0.77		0.73		0.53		0.67		0.71		0.77	
SIMVERB	0.43	0.27	0.36		0.23		0.37		0.15		0.19		0.53	
WORDSIM353	0.58	0.61	0.70		0.61		0.47		0.67		0.59		0.72	
SIMLEX999-N	0.44	0.33	0.45		0.39		0.48		0.32		0.34		0.68	
MEN-N	0.72	0.68	0.77	___	0.76	___	0.57	___	0.71	___	0.73	___	0.78	___

Sense embeddings using retrofitting

Unsupervised WSD SemEval'13, **ReprL4NLP** [Pelevina et al., 2016]:

- comparable to SOTA, incl. sense embeddings.

Sense embeddings using retrofitting

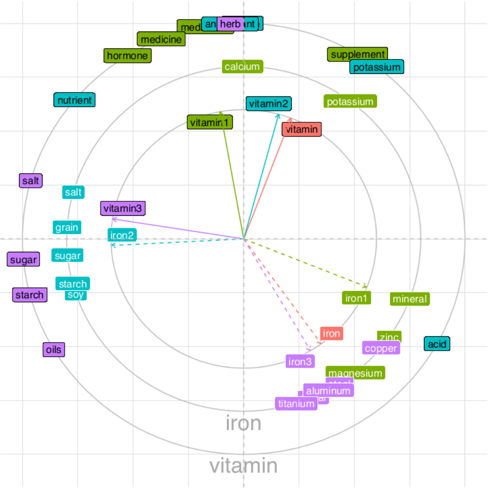
Unsupervised WSD SemEval'13, **ReprL4NLP** [Pelevina et al., 2016]:

- comparable to SOTA, incl. sense embeddings.

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

	AUTOEXTEND	ADAGRAM	SGNS	SGNS+SENSES	GLOVE	GLOVE+SENSES	SYMPAT	SYMPAT+SENSES	LSABOW	LSABOW+SENSES	LSAHAL	LSAHAL+SENSES	PARAGRAMSL	PARAGRAMSL+SENSES
SIMLEX999	0.45	0.29	0.44	**0.46**	0.37	**0.41**	0.54	**0.55**	0.30	**0.39**	0.27	**0.38**	**0.68**	0.64
MEN	0.72	0.67	0.77	**0.78**	0.73	**0.77**	0.53	**0.68**	0.67	**0.70**	0.71	**0.74**	0.77	**0.80**
SIMVERB	0.43	0.27	0.36	**0.39**	0.23	**0.30**	0.37	**0.45**	0.15	**0.22**	0.19	**0.28**	0.53	0.53
WORDSIM353	0.58	0.61	**0.70**	0.69	0.61	**0.65**	0.47	**0.62**	**0.67**	0.66	0.59	**0.63**	0.72	**0.73**
SIMLEX999-N	0.44	0.33	0.45	**0.50**	0.39	**0.47**	0.48	**0.55**	0.32	**0.46**	0.34	**0.44**	**0.68**	0.66
MEN-N	0.72	0.68	0.77	**0.79**	0.76	**0.80**	0.57	**0.74**	0.71	**0.73**	0.73	**0.76**	0.78	**0.81**

Sense embeddings using retrofitting



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

LREC'18 [Remus & Biemann, 2018]

Watset: synset induction

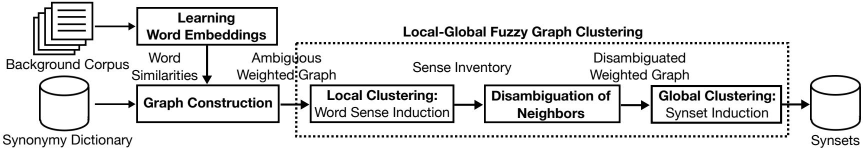
ACL'17 [Ustalov et al., 2017b]

Examples of extracted synsets:

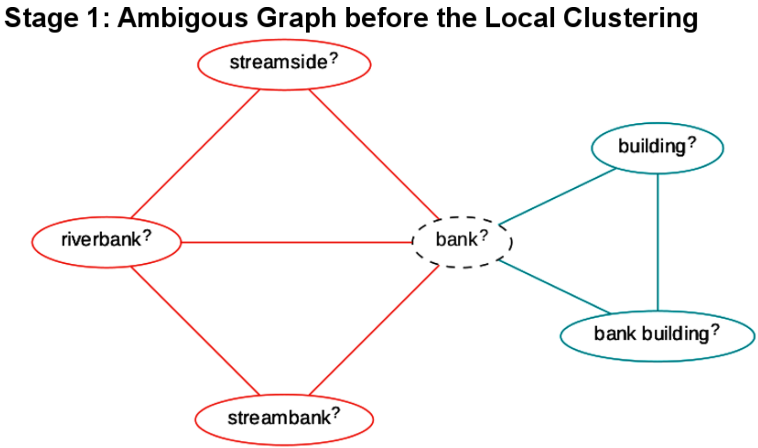
Size	Synset
2	{*decimal point, dot*}
3	{*gullet, throat, food pipe*}
4	{*microwave meal, ready meal, TV dinner, frozen dinner*}
5	{*objective case, accusative case, oblique case, object case, accusative*}
6	{*radio theater, dramatized audiobook, audio theater, radio play, radio drama, audio play*}

Synset induction

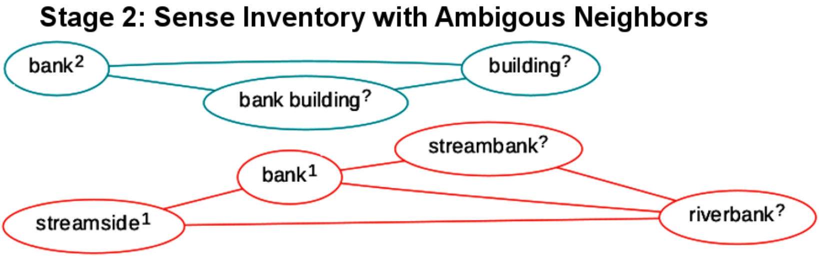
Outline of the 'Watset' method:



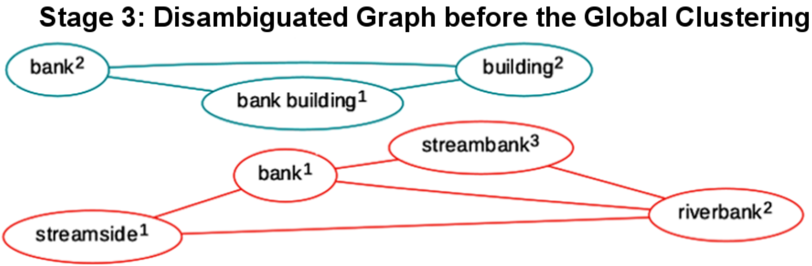
Synset induction



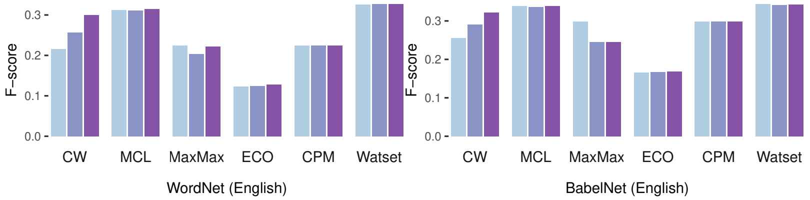
Synset induction



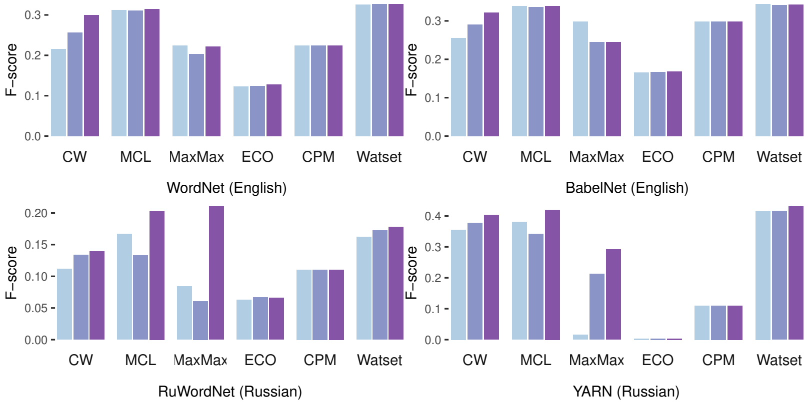
Synset induction



Synset induction



Synset induction

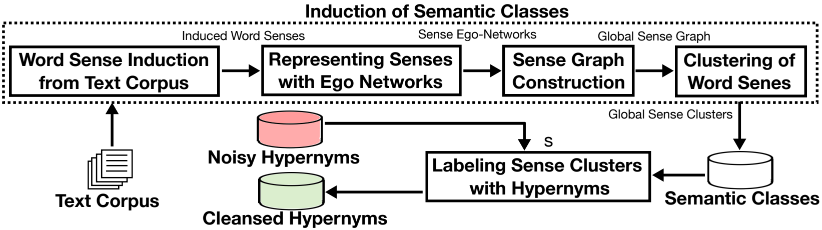


Induction of semantic classes

Examples of semantic classes:

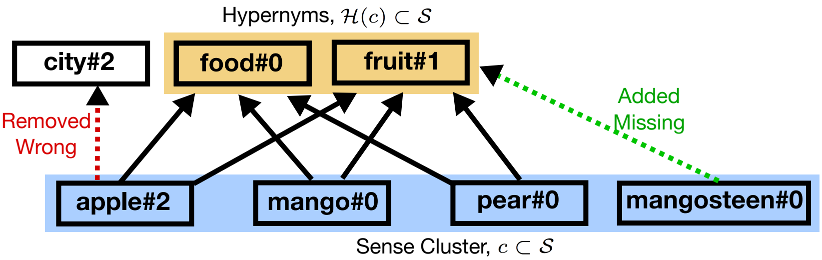
ID	Sense Cluster	Hypernyms
1	peach#1, banana#1, pineapple#0, berry#0, blackberry#0, grapefruit#0, strawberry#0, blueberry#0, fruit#0, grape#0, melon#0, orange#0, pear#0, plum#0, raspberry#0, watermelon#0, apple#0, apricot#0, ...	vegetable#0, fruit#0, crop#0, ingredient#0, food#0, ·
2	C#4, Basic#2, Haskell#5, 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, ...	programming language#3, technology#0, language#0, format#2, app#0

Induction of semantic classes



Induction of sense semantic classes

Filtering noisy hypernyms with semantic classes
LREC'18 [Panchenko et al., 2018]:



Induction of sense semantic classes

Filtering of a noisy hypernymy database with semantic classes.
LREC'18 [Panchenko et al., 2018]

	Precision	**Recall**	**F-score**
Original Hypernyms (Seitner et al., 2016)	0.475	0.546	0.508
Semantic Classes (coarse-grained)	**0.541**	**0.679**	**0.602**

Making induced senses interpretable

Making induced senses interpretable

Knowledge-based sense representations are **interpretable**

• bn:01713224n • NOUN • Named Entity • Categories: High-level programming languages, Dutch inventions, Class-based programming languages, Cross platform free software...

EN **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

IS A	programming language • free software • scripting language
HAS PART	pandas
HAS KIND	Stackless Python
DESIGNER	Guido van Rossum
DEVELOPER	Python Software Foundation • Guido van Rossum
DIALECTS	Cython • Stackless Python
INFLUENCED BY	ALGOL 68 • alphabet • ruby
LICENSE	Python Software Foundation License

More relations

EXPLORE NETWORK

Making induced senses interpretable

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, -0.0118133 ,  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

Sentence
Jaguar is a large spotted predator of tropical America similar to the leopard. A

Word
Jaguar B

Model
Word Senses based on Cluster Word Features C

http://jobimtext.org/wsd

PREDICT SENSE | RANDOM SAMPLE

Predicted senses for 'Jaguar'

1. jaguar (animal)

Similarity score: 0.00184 / Confidence: 99.87% / Sense ID: jaguar#0 / BabelNet ID: bn:00033987n

Hypernyms: animal, wildlife, bird, mammal D

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.

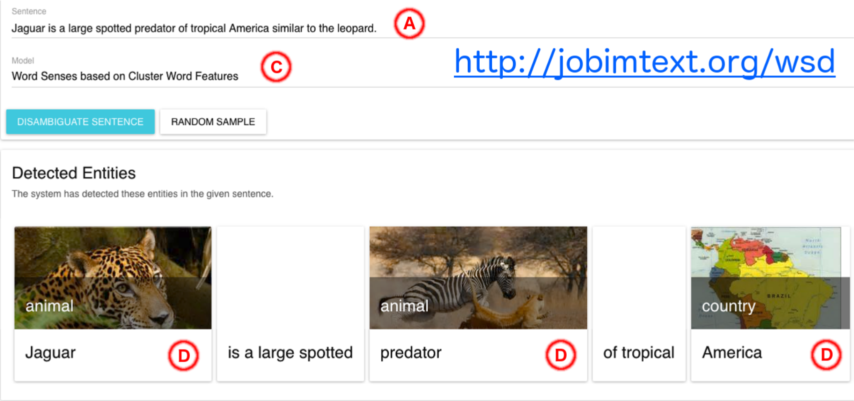
Cluster words: lion, tiger, leopard, wolf, monkey, otter, crocodile, alligator, deer, cat, elephant, fox, eagle, owl, snake

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

BABELNET LINK F SHOW LESS E

Making induced senses interpretable



Hypernymy prediction in context. **EMNLP’17** [Panchenko et al., 2017b]

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Making induced senses interpretable

- **11.702 sentences**, **863 words** with **avg.polysemy of 3.1**.

WSD Model		**Accuracy**	
Inventory	Features	Hypers	HyperHypers
Word Senses	Random	0.257	0.610
Word Senses	MFS	0.292	0.682
Word Senses	Cluster Words	0.291	0.650
Word Senses	Context Words	**0.308**	**0.686**

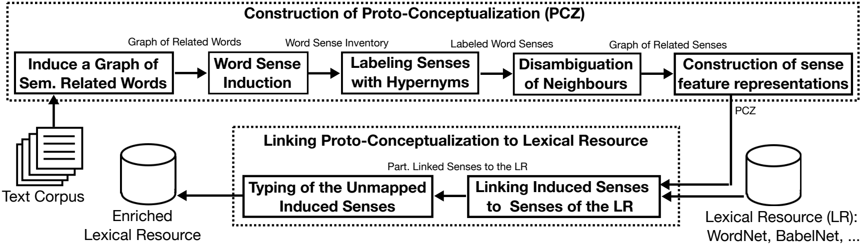
Making induced senses interpretable

- **11.702 sentences**, **863 words** with **avg.polysemy of 3.1**.

WSD Model		**Accuracy**	
Inventory	Features	Hypers	HyperHypers
Word Senses	Random	0.257	0.610
Word Senses	MFS	0.292	0.682
Word Senses	Cluster Words	0.291	0.650
Word Senses	Context Words	**0.308**	**0.686**
Super Senses	Random	0.001	0.001
Super Senses	MFS	0.001	0.001
Super Senses	Cluster Words	**0.174**	**0.365**
Super Senses	Context Words	0.086	0.188

Linking induced senses to resources

Linking induced senses to resources



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

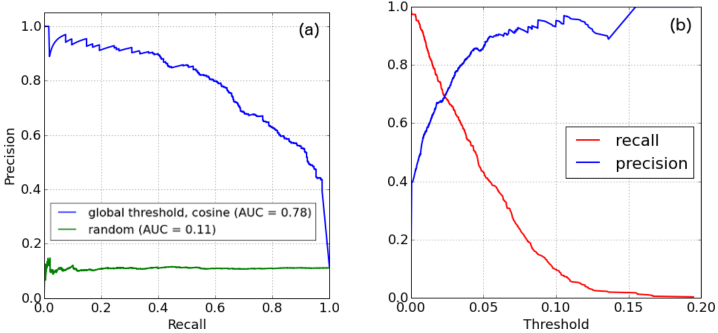
Word	AdaGram	BabelNet	AdaGram BoW	BabelNet BoW
python	2	bn:01713224n	perl, php, java, smalltalk, ruby, lua, tcl, scripting, javascript, bindings, binding, programming, coldfusion, actionscript, net, . . .	language, programming, pythonista, python programming, python3, python2, level, computer, pythonistas, python3000,
python	1	bn:01157670n	monty, circus, spamalot, python, magoo, muppet, snoopy, featurette, disney, tunes, tune, classic, shorts, short, apocalypse, . . .	monty, comedy, monty python, british, monte, monte python, troupe, pythonesque, foot, artist, record, surreal, terry, . . .
python	3	bn:00046456n	spectacled, unicornis, snake, giant, caiman, leopard, squirrel, crocodile, horned, cat, mole, elephant, opossum, pheasant, . . .	molurus, indian, boa, tigris, tiger python, rock, tiger, indian python, reptile, python molurus, indian rock python, coluber, . . .
python	4	bn:01157670n	circus, fly, flying, dusk, lizard, moth, unicorn, puff, adder, vulture, tyrannosaurus, zephyr, badger, . . .	monty, comedy, monty python, british, monte, monte python, troupe, pythonesque, foot, artist, record, surreal, terry, . . .
python	1	bn:00473212n	monty, circus, spamalot, python, magoo, muppet, snoopy, featurette, disney, tunes, tune, classic, shorts, short, apocalypse, . . .	pictures, monty, python monty pictures, limited, company, python pictures limited, kingdom, picture, serve, director, . . .
python	1	bn:03489893n	monty, circus, spamalot, python, magoo, muppet, snoopy, featurette, disney, tunes, tune, classic, shorts, short, apocalypse, . . .	film, horror, movie, clabaugh, richard, monster, century, direct, snake, python movie, television, giant, natural, language, for-tv, . . .

Linking induced senses to resources

Model	Representation of the Sense "disk (medium)"
WordNet	memory, device, floppy, disk, hard, disk, disk, computer, science, computing, diskette, fixed, disk, floppy, magnetic, disc, magnetic, disk, hard, disc, storage, device
WordNet + Linked	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, modem, 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

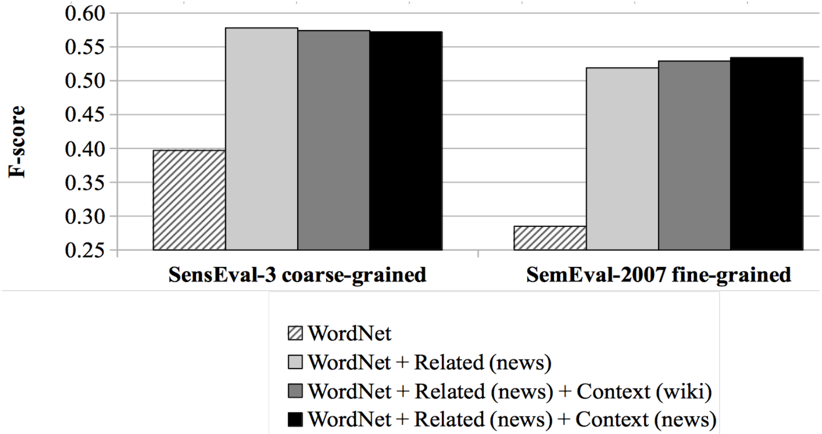
Linking induced senses to resources

Evaluation of linking accuracy:



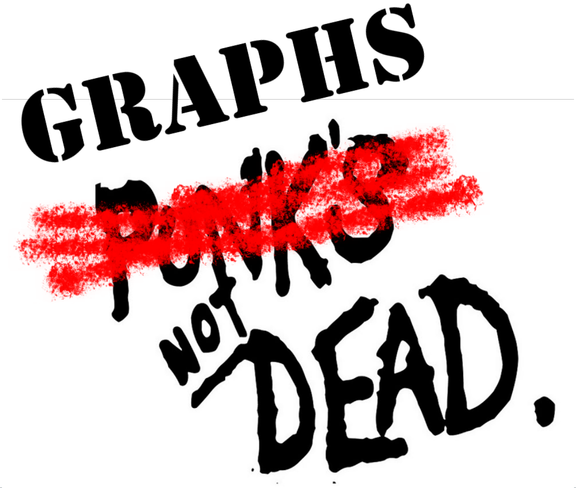
Linking induced senses to resources

Evaluation of enriched representations based on WSD:



Conclusion

Vectors + Graphs = ♡



Take home messages

- We can **induce word senses**, **synsets** and **semantic classes** in a knowledge-free way using **graph clustering** and **distributional models**.

Take home messages

- 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**.

Take home messages

- 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!
- **More details**: `http://russe.nlpub.org/2018/wsi`

Acknowledgments

Thank you! Questions?

This research was supported by

Deutsche Forschungsgemeinschaft

Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

Sense embeddings using retrofitting

Evaluation on SemEval 2013 Task 13 WSI&D:

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

Bartunov, S., Kondrashkin, D., Osokin, A., & Vetrov, D. (2016).
Breaking sticks and ambiguities with adaptive skip-gram.
In *Artificial Intelligence and Statistics* (pp. 130–138).

Biemann, C., Faralli, S., Panchenko, A., & Ponzetto, S. P. (2018).
A framework for enriching lexical semantic resources with distributional semantics.
In *Journal of Natural Language Engineering* (pp. 56–64).: Cambridge Press.

Faralli, S., Panchenko, A., Biemann, C., & Ponzetto, S. P. (2016).
Linked disambiguated distributional semantic networks.
In *International Semantic Web Conference* (pp. 56–64).: Springer.

Panchenko, A. (2016).
Best of both worlds: Making word sense embeddings interpretable.
In *LREC*.

Panchenko, A., Faralli, S., Ponzetto, S. P., & Biemann, C. (2017a).
Using linked disambiguated distributional networks for word sense disambiguation.
In *Proceedings of the 1st Workshop on Sense, Concept and Entity Representations and their Applications* (pp. 72–78). Valencia, Spain: Association for Computational Linguistics.

Panchenko, A., Marten, F., Ruppert, E., Faralli, S., Ustalov, D., Ponzetto, S. P., & Biemann, C. (2017b).
Unsupervised, knowledge-free, and interpretable word sense disambiguation.
In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 91–96). Copenhagen, Denmark: Association for Computational Linguistics.

Panchenko, A., Ruppert, E., Faralli, S., Ponzetto, S. P., & Biemann, C. (2017c).

Unsupervised does not mean uninterpretable: The case for word sense induction and disambiguation.
In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers* (pp. 86–98). Valencia, Spain: Association for Computational Linguistics.

Panchenko, A., Ustalov, D., Faralli, S., Ponzetto, S. P., & Biemann, C. (2018).
Improving hypernymy extraction with distributional semantic classes.
In *Proceedings of the LREC 2018* Miyazaki, Japan: European Language Resources Association.

Pelevina, M., Arefiev, N., Biemann, C., & Panchenko, A. (2016).
Making sense of word embeddings.
In *Proceedings of the 1st Workshop on Representation Learning for NLP* (pp. 174–183). Berlin, Germany: Association for Computational Linguistics.

Remus, S. & Biemann, C. (2018).

Retrofittingword representations for unsupervised sense aware word similarities.
In *Proceedings of the LREC 2018* Miyazaki, Japan: European Language Resources Association.

Rothe, S. & Schütze, H. (2015).
Autoextend: Extending word embeddings to embeddings for synsets and lexemes.
In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 1793–1803). Beijing, China: Association for Computational Linguistics.

Ustalov, D., Chernoskutov, M., Biemann, C., & Panchenko, A. (2017a).
Fighting with the sparsity of synonymy dictionaries for automatic synset induction.
In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 94–105).: Springer.

- Ustalov, D., Panchenko, A., & Biemann, C. (2017b).
  Watset: Automatic induction of synsets from a graph of synonyms.
  In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1579–1590). Vancouver, Canada: Association for Computational Linguistics.

- Ustalov, D., Teslenko, D., Panchenko, A., Chernoskutov, M., & Biemann, C. (2018).
  Word sense disambiguation based on automatically induced synsets.
  In *LREC 2018, 11th International Conference on Language Resources and Evaluation : 7-12 May 2018, Miyazaki (Japan)* (pp. tba). Paris: European Language Resources Association, ELRA-ELDA.
  Accepted for publication.