



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Alexander Panchenko

**INDUCING INTERPRETABLE WORD
SENSES FOR WSD AND ENRICHMENT OF
LEXICAL RESOURCES**

Sparse sense representations



Sparse sense representations



Watset: synset induction

Watset: synset induction



Watset: synset induction



Induction of sense semantic classes



Induction of sense semantic classes



Induction of sense semantic classes

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...

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:00033967n

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

Making induced senses interpretable

Linking induced senses to resources



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Linking induced senses to resources



Linking induced senses to resources

Linking induced senses to resources

Linking induced senses to resources

Linking induced senses to resources



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Linking induced senses to resources



Linking induced senses to resources