

Interpretable Natural Language Processing

Why and How?



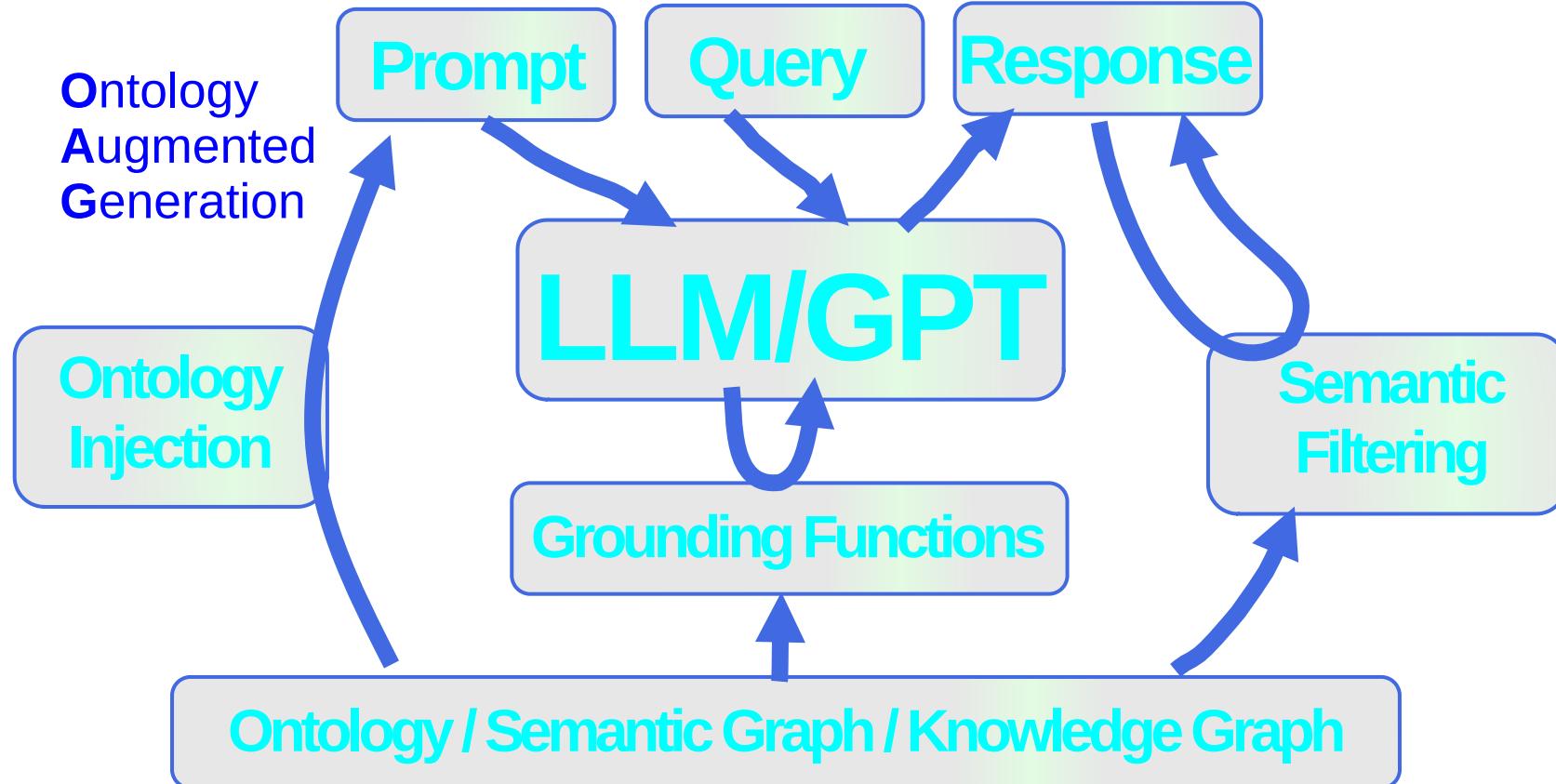
<https://agirussia.org>

Anton Kolonin
akolonin@aigents.com
Messengers: akolonin



Why do we need INLP in LLM/DL/GPT era?

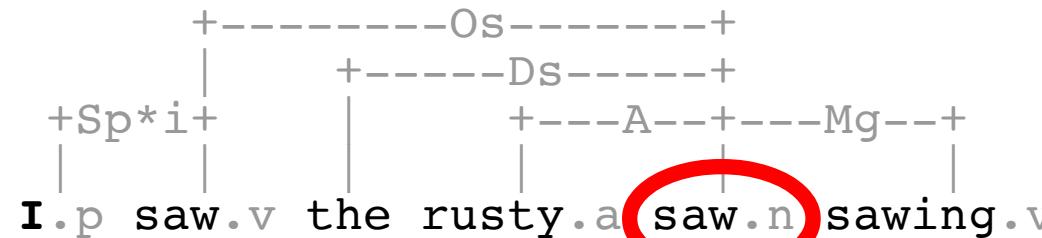
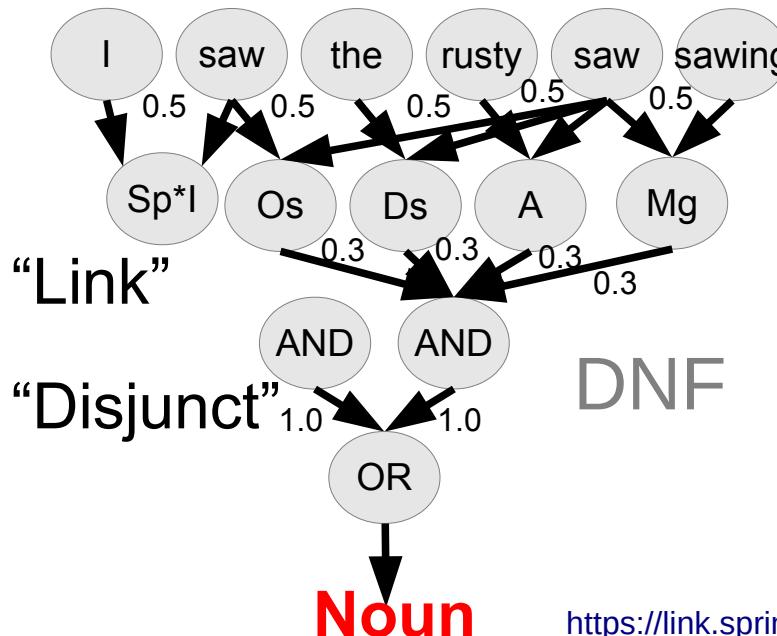
Fighting Hallucinations with OAG/RAG at Palantir



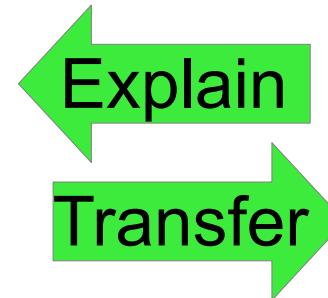
<https://blog.palantir.com/reducing-hallucinations-with-the-ontology-in-palantir-aip-288552477383>

Bridging the Symbolic-Subsymbolic gap in NLP

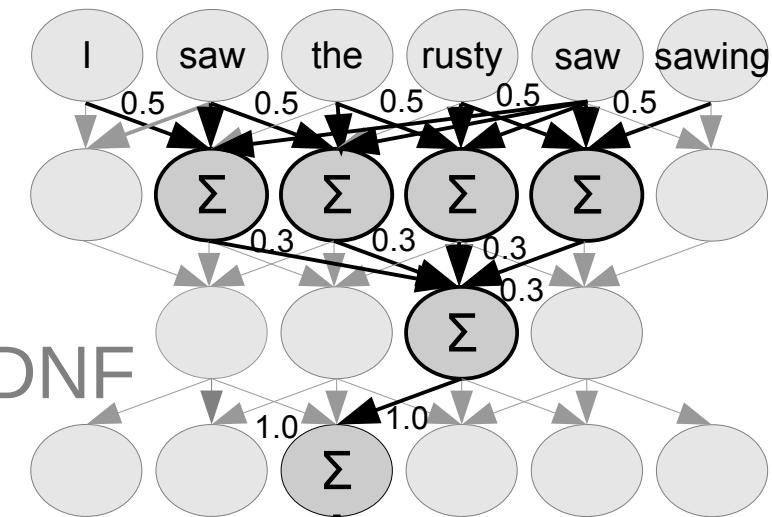
Formal
Link Grammar



Deep
Language Model



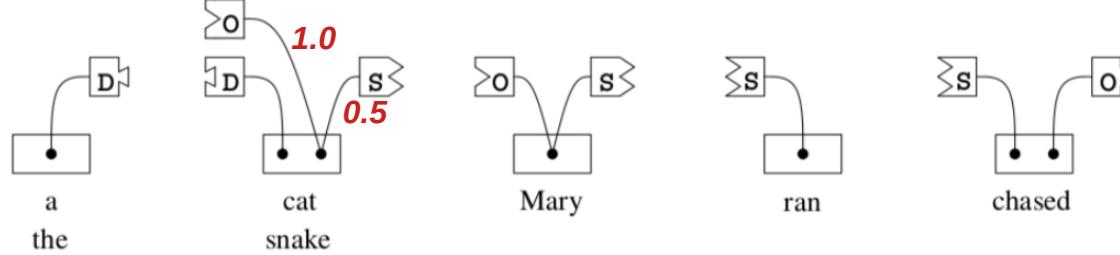
Soft-DNF



https://link.springer.com/chapter/10.1007/978-3-319-97676-1_11
https://doi.org/10.1007/978-3-030-27005-6_11
<https://github.com/singnet/language-learning/>

Copyright © 2024 Anton Kolonin, Aigents®

Link Grammar – Connectors/Costs, Disjuncts

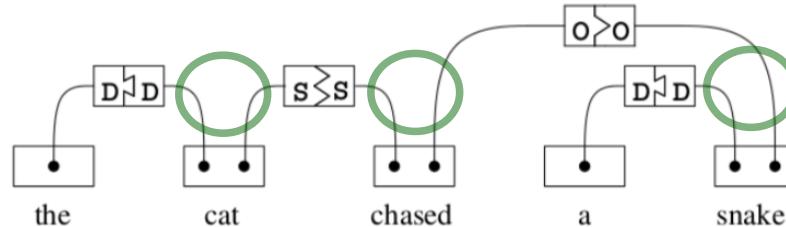


Connectors & Costs

An illustration of Link Grammar connectors and disjuncts. The connectors are the jigsaw-puzzle-shaped pieces; connectors are allowed to connect only when the tabs fit together. A disjunct is the entire (ordered) set of connectors for a word. As lexical entries appearing in a dictionary, the above would be written as

```
a the: D+;  
cat snake: D- & (S+ or O-);  
Mary: O- or S+;  
ran: S-;  
chased S- & O+;
```

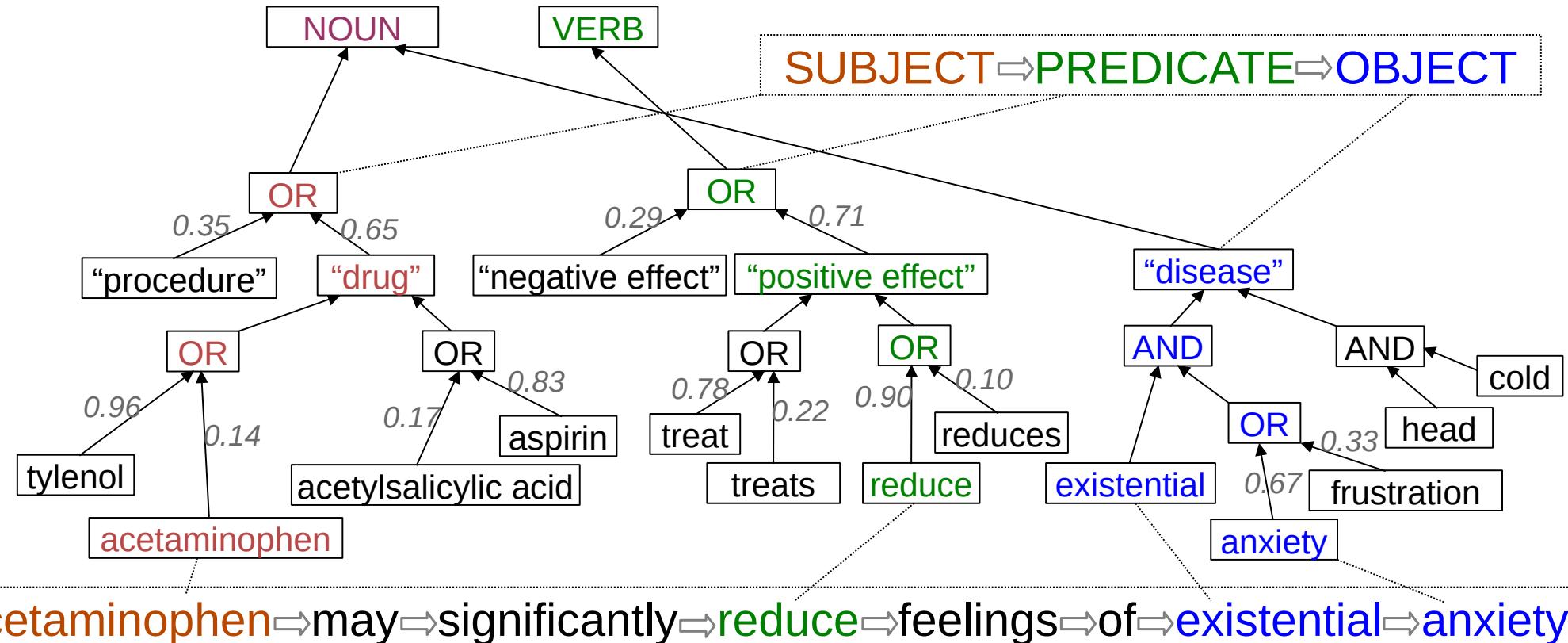
Note that although the symbols ‘‘&’’ and ‘‘or’’ are used to write down disjuncts, these are **not** Boolean operators, and do **not** form a Boolean algebra. They do form a non-symmetric compact closed monoidal algebra. The diagram below illustrates puzzle pieces, assembled to form a parse:



Disjuncts of Connectors

<https://arxiv.org/abs/1401.3372>
https://en.wikipedia.org/wiki/Link_grammar
<https://github.com/opencog/link-grammar>

NLP patterns (words, punctuation, phrases) for Interpretable Natural Language Processing (Aigents® “Deep Patterns”)



<https://ieeexplore.ieee.org/document/7361868>
<https://github.com/aigents/aigents-java>

Generalized Text Mining with Aigents® “Deep Patterns”

Classification

Category:
“Healthcare”

tylenol
acetaminophen
placebo

Case/Relationship Extraction

Entity (Case): “Treatment:
Healing anxiety with Tylenol”

significantly
reduce
feelings
study

Here's the Tylenol twist: Before they began writing, half of each group received acetaminophen while the other half swallowed a placebo. Even among those people who wrote about death, the Tylenol takers set bail at roughly \$300—a sign that acetaminophen may significantly reduce feelings of existential anxiety, explains study lead author Daniel Randles, a PhD candidate in UBC's department of... psychology.

<https://ieeexplore.ieee.org/document/7361868>
<https://github.com/aigents/aigents-java>

Property Attribution Entity Extraction

Brand: Tylenol
Substance: acetaminophen
Reliability: medium
Effect: positive
Diagnosis: Anxiety
Reporter: Daniel Randles

acetaminophen
may
reduce
anxiety
explains

acetaminophen may
significantly reduce
feelings of existential
anxiety, explains study
lead author Daniel
Randles.

Case 1: Property Attribution / Aigents® “Deep Patterns”

```
<set> := <disjunctive-set> | <conjunctive-set> | <M-skip-N-gram>
<disjunctive-set> := { <pattern> * }
<conjunctive-set> := ( <pattern> * )
<N-gram> := [ <pattern> * ]
<pattern> := <token> | <regexp> | <variable> | <set>
```

Variables may have domain restrictions
in ontology and/or refer to other
patterns as subgraphs

Example:

```
{[$description catheter] [$coating coating] [$inner-diameter
    {diameter inner-diameter}] [$tip tip] [$pattern pattern]}
```

X

“Convey Guiding Catheter. Unique hydrophilic coating.

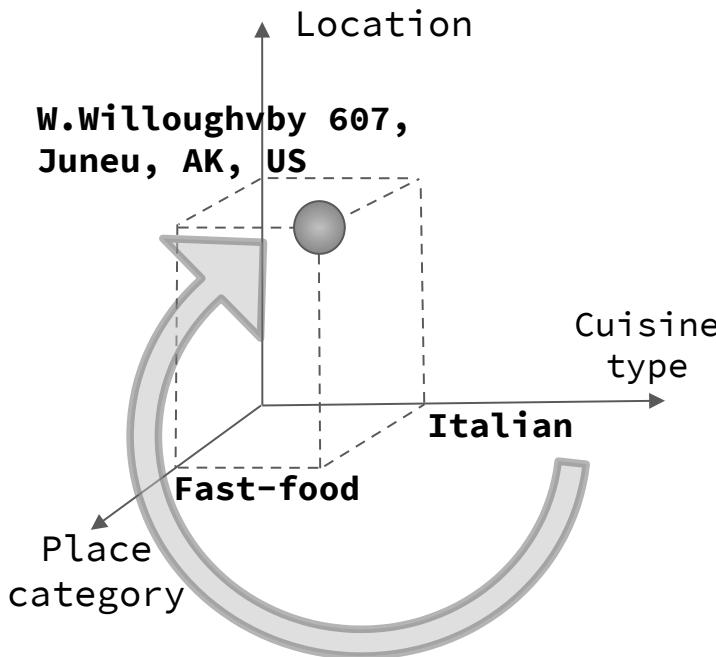
 Smallatraumatic soft tip. Ultra-thin 1 × 2 flat wire braid pattern”

=

```
{ coating : "hydrophilic", description : "convey guiding",
    pattern : "ultra-thin 1 × 2 flat wire braid", tip : "soft" }
```

<https://ieeexplore.ieee.org/document/7361868>
<https://github.com/aigents/aigents-java>

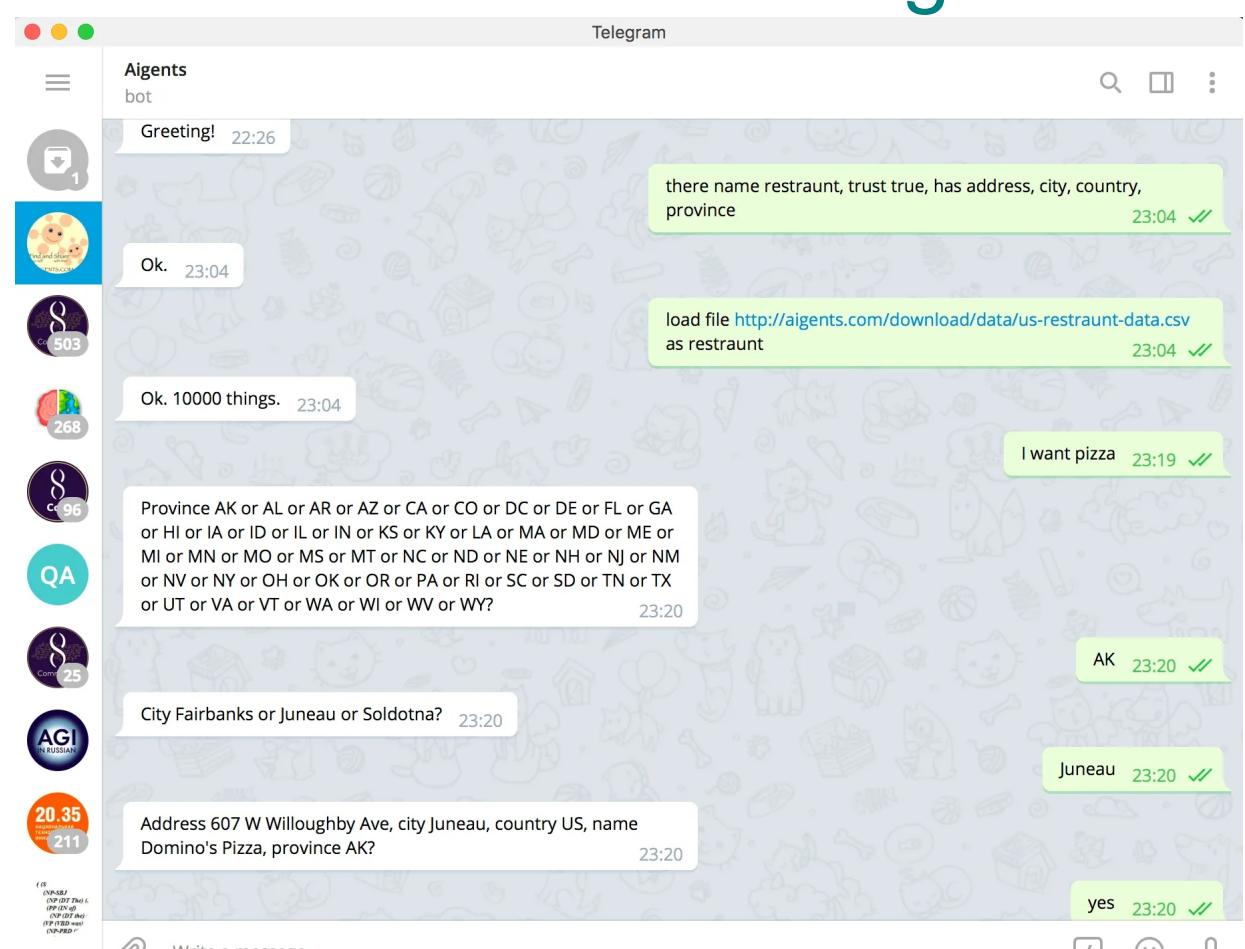
Case 8: Data-driven Conversational Intelligence



Dynamic data-driven dialog flow minimizing uncertainty
(Can be used for OAG/TAG)

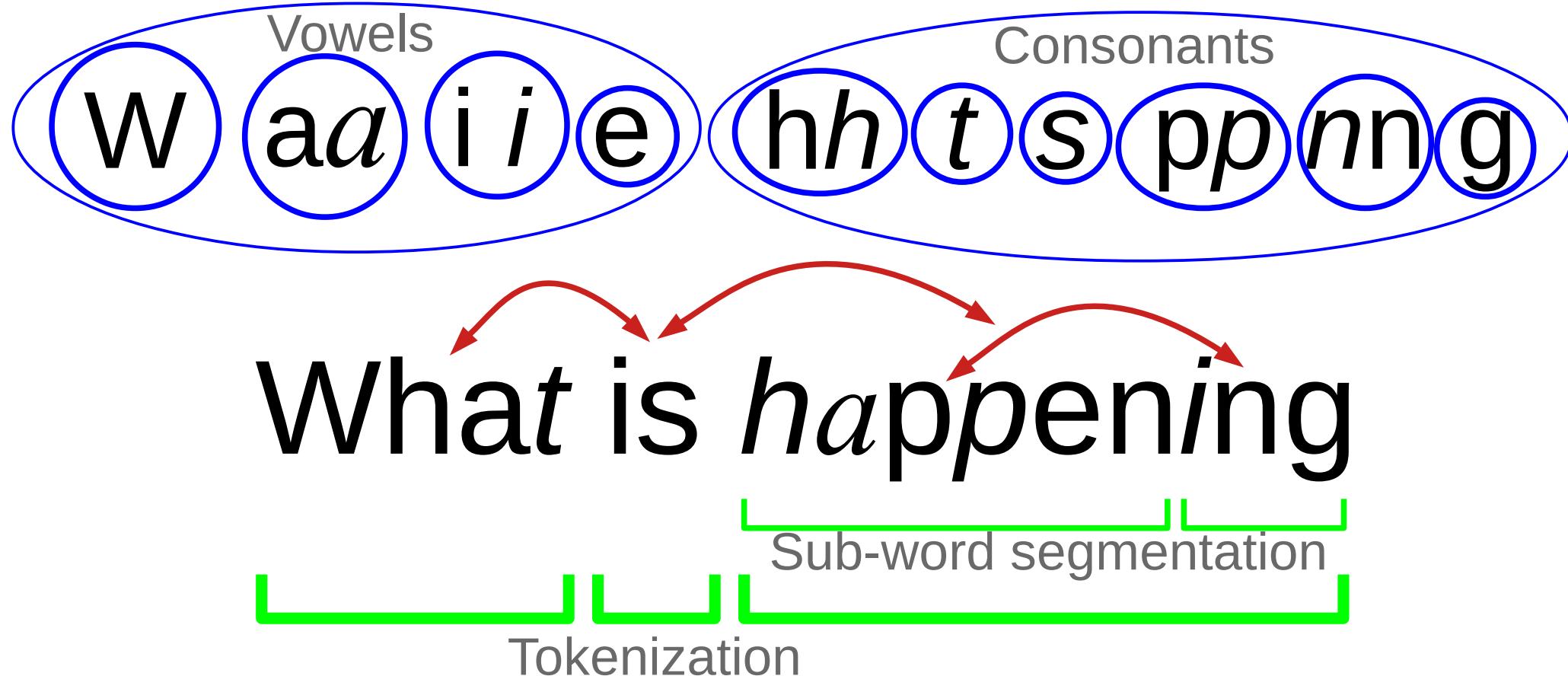
<https://t.me/AigentsBot>
<https://blog.singularitynet.io/chat-with-your-data-using-aigents-bots-99b76cae65f2>

Copyright © 2024 Anton Kolonin, Aigents®

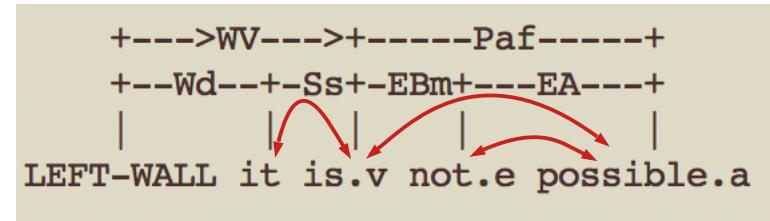
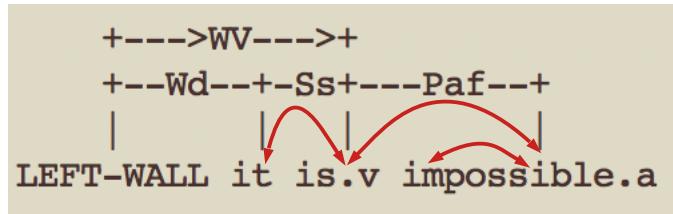


Learning Lexicon, Punctuation, Morphology and Grammar

Clustering, Segmentation and Parsing



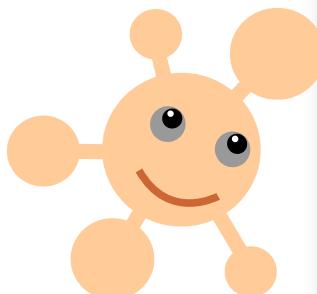
Blurring boundaries between grammatical parsing and morphological parsing causes tokenization ambiguity



这不可能
(Google)

这 不 可 能
(Jieba)

Minimizing Uncertainty



New Tab × +

how are you

- how are you - Google Search
- how are
- how are you doing
- how are you answers
- How Are You Feeling - Song by TAYLOR DEE
- How Are You Today? - Song by Maple Leaf Learning
- how are you doing answer
- how are you synonyms
- how are you in spanish
- how are things going



New Tab × +

how many

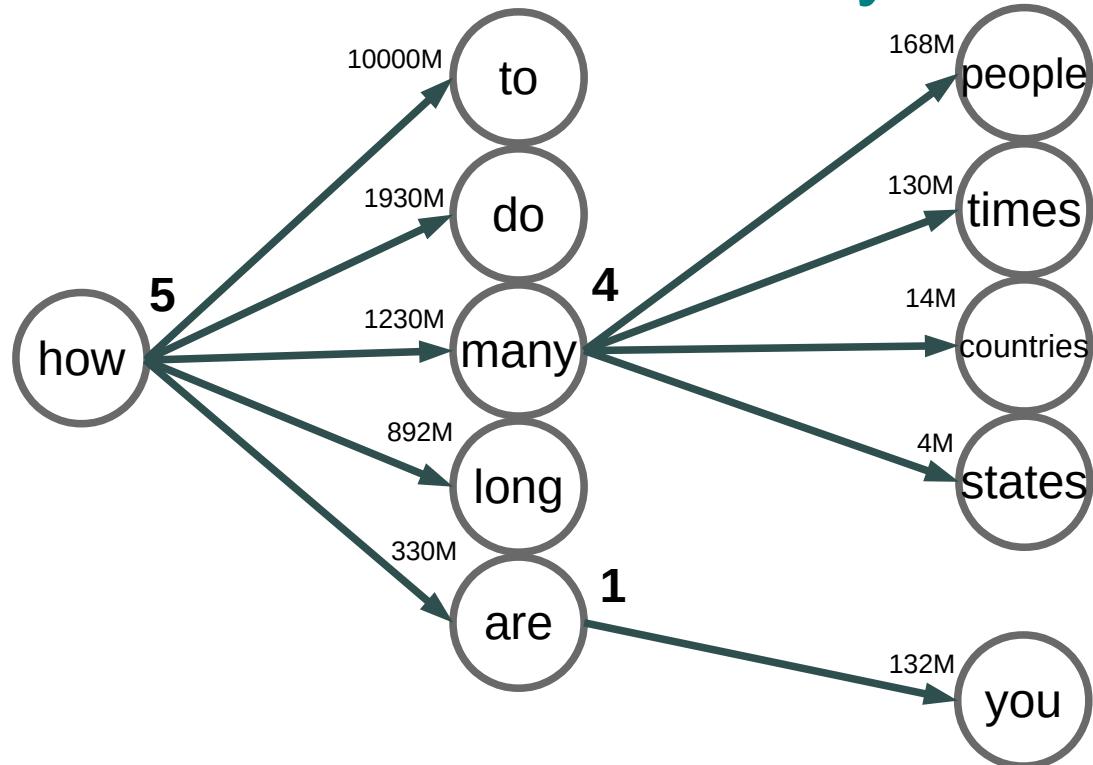
- how many - Google Search
- how many countries in the world
- how many weeks in a year
- how many states in usa
- how many continents
- how many people in the world
- how many words
- how many continents are there
- how many bones in human body
- how many episodes in house of dragons

<https://aclanthology.org/2022.emnlp-main.239/>

<https://arxiv.org/pdf/2303.02427.pdf>

Copyright © 2024 Anton Kolonin, Aigents®

Unsupervised Learning for Text Segmentation based on Probability and Uncertainty Measures



Metrics/Indicators:

Mutual Information¹
Conditional Probability^{1,2}
Transition Freedom^{2,3}

¹ <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=6983&context=etd>

² <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2655800/>

³ Karl Friston. The free-energy principle: a unified brain theory? <https://www.nature.com/articles/nrn2787>

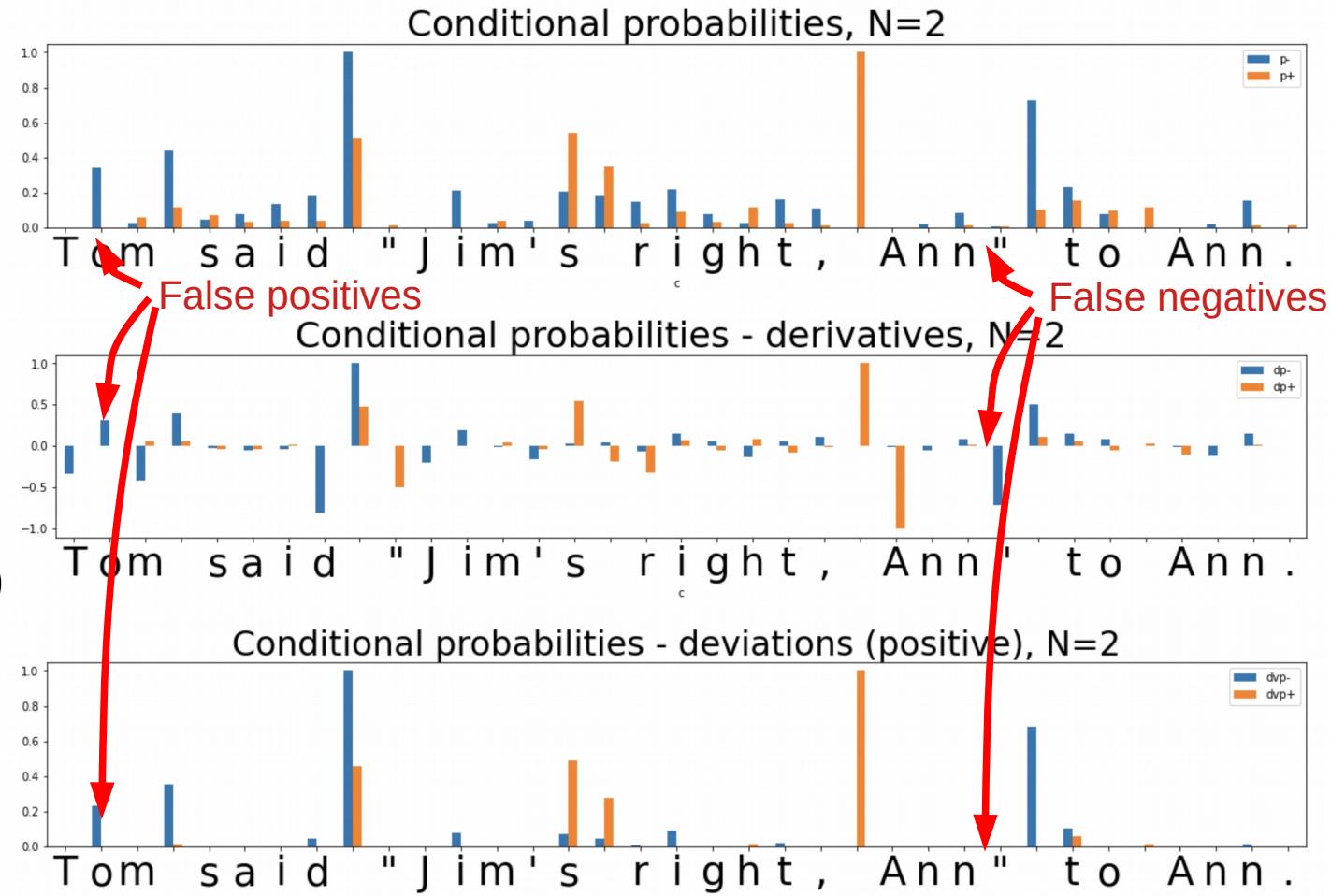
Case 5: Unsupervised Text Segmentation (Probability)

Metrics/Indicators:

Ngram (Character)
Conditional
Probability
(of Transition)

$P(\text{Ngram}_{n+1})/P(\text{Ngram}_n)$

$P("m")/P(m")$



<https://aclanthology.org/2022.emnlp-main.239/>

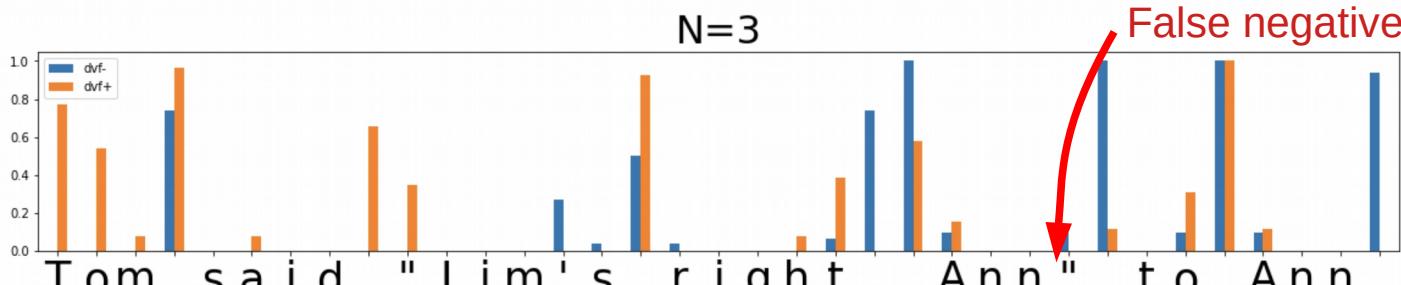
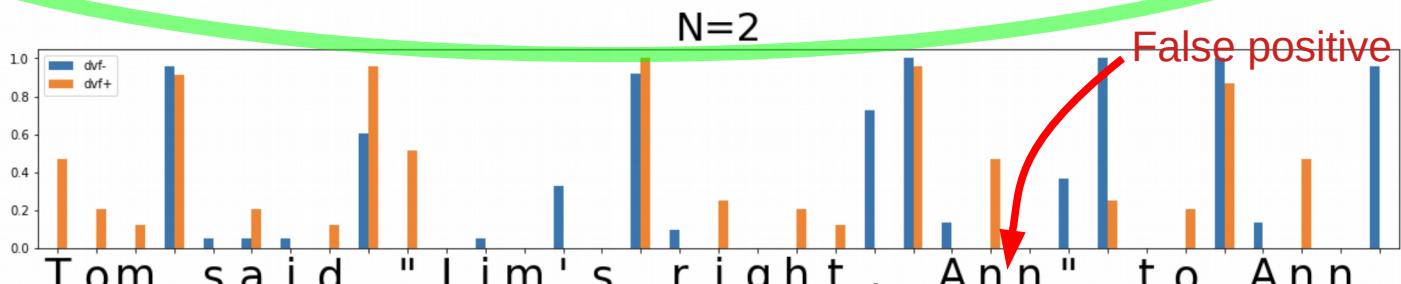
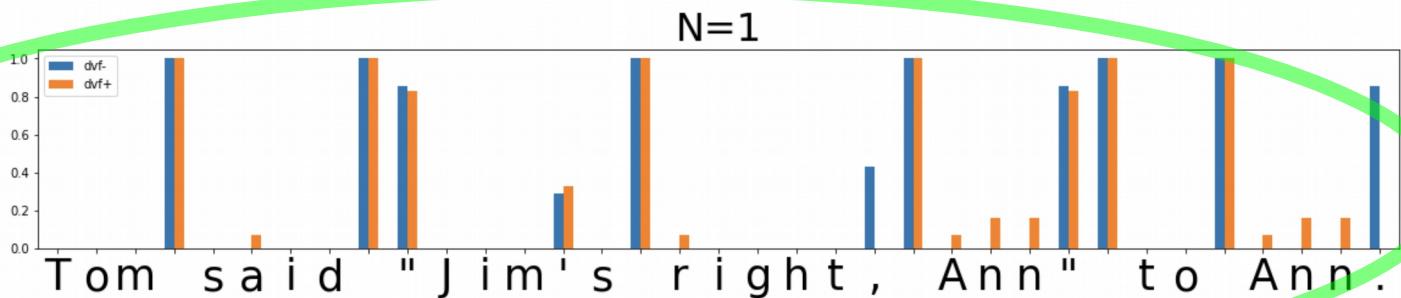
Copyright © 2024 Anton Kolonin, Agents®

Case 5: Unsupervised Text Segmentation (Uncertainty)

Metrics/
Indicators:

Transition
Freedom
Deviation

(varying “N”
of N-gram)



Case 5: Unsupervised Text Segmentation ("Freedom"-based Tokenization against Lexicon-based one, referring to Rule-based)

Language	Tokenizer	Tokenization F_1	Lexicon Discovery Precision
English	Freedom-based	0.99	0.99 (vs. 1.0)
English	Lexicon-based*	0.99	-
Russian	Freedom-based	1.0	1.0 (vs. 1.0)
Russian	Lexicon-based*	0.94	-
Chinese	Freedom-based	0.71	0.92 (vs. 0.94)
Chinese	Lexicon-based*	0.83	-

*Lexicon-based Tokenization - greedy/beam search on word length (optimal) or frequency

<https://aclanthology.org/2022.emnlp-main.239/>

<https://ieeexplore.ieee.org/document/10347856>

https://link.springer.com/chapter/10.1007/978-3-031-44865-2_1

Case 5: Unsupervised Text Segmentation (Hyper-parameters)

English

F1 - Brown ddf- & ddf+ filter=0 parameters=10967135

Hyper-Parameters:

Metric:
Transition
Freedom

[1]	0.5	0.75	0.82	0.79	0.79	0.81	0.89	0.89	0.89
[2]	0.46	0.54	0.62	0.67	0.85	0.92	0.81	0.71	0.37
[3]	0.56	0.67	0.72	0.73	0.69	0.61	0.46	0.36	0.19
[4]	0.54	0.68	0.7	0.6	0.43	0.3	0.19	0.15	0.1
[5]	0.51	0.55	0.52	0.38	0.25	0.16	0.11	0.1	0.08
[6]	0.48	0.46	0.38	0.25	0.17	0.12	0.1	0.08	0.07
[7]	0.42	0.34	0.24	0.15	0.11	0.1	0.08	0.08	0.07
[1, 2]	0.47	0.58	0.82	0.94	0.94	0.91	0.89	0.79	0.56
[2, 3]	0.51	0.62	0.74	0.79	0.83	0.81	0.66	0.46	0.24
[1, 2, 3]	0.5	0.69	0.79	0.87	0.91	0.89	0.78	0.58	0.25
[1, 2, 3, 4]	0.55	0.75	0.84	0.86	0.84	0.75	0.52	0.31	0.15
[4, 5, 6, 7]	0.56	0.6	0.51	0.33	0.2	0.14	0.1	0.08	0.07
[1, 2, 3, 4, 5]	0.56	0.78	0.86	0.84	0.74	0.53	0.31	0.17	0.1
[1, 2, 3, 4, 5, 6, 7]	0.59	0.78	0.82	0.69	0.49	0.26	0.15	0.09	0.07
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

Threshold
for model
compression

F1 - Brown ddf- & ddf+ filter=0.0001 parameters=8643703

Combination
of Ngram N-s

[1]	0.73	0.96	0.98	0.99	0.96	0.94	0.95	0.95	0.89
[2]	0.46	0.54	0.64	0.77	0.91	0.94	0.89	0.7	0.44
[3]	0.55	0.66	0.74	0.78	0.72	0.65	0.49	0.37	0.19
[4]	0.54	0.67	0.7	0.61	0.45	0.32	0.21	0.16	0.1
[5]	0.51	0.55	0.52	0.38	0.26	0.17	0.12	0.1	0.08
[6]	0.48	0.46	0.38	0.26	0.18	0.13	0.1	0.09	0.07
[7]	0.42	0.35	0.25	0.16	0.12	0.1	0.09	0.08	0.08
[1, 2]	0.51	0.64	0.82	0.96	0.96	0.96	0.9	0.88	0.68
[2, 3]	0.5	0.62	0.74	0.85	0.89	0.86	0.71	0.51	0.27
[1, 2, 3]	0.53	0.69	0.81	0.91	0.93	0.92	0.82	0.6	0.36
[1, 2, 3, 4]	0.55	0.75	0.86	0.88	0.88	0.81	0.57	0.33	0.17
[4, 5, 6, 7]	0.56	0.6	0.52	0.35	0.22	0.15	0.1	0.09	0.07
[1, 2, 3, 4, 5]	0.57	0.79	0.88	0.86	0.78	0.59	0.33	0.18	0.1
[1, 2, 3, 4, 5, 6, 7]	0.59	0.79	0.83	0.71	0.5	0.28	0.16	0.09	0.08
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

Threshold for
segmentation

Case 5: Unsupervised Text Segmentation (Auto-learning)

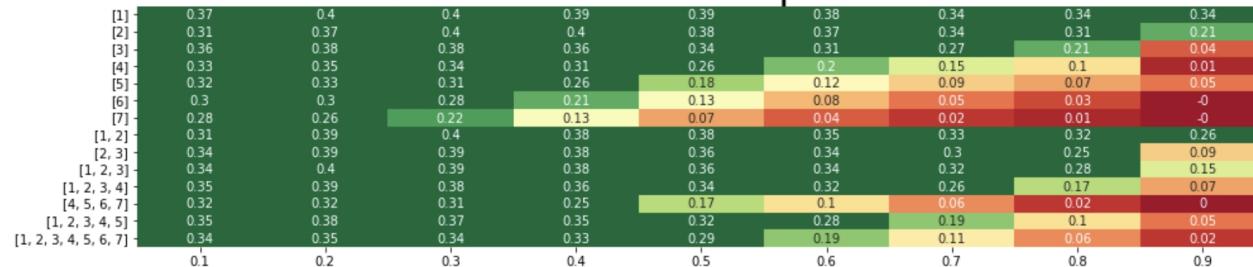
English,
Brown corpus,
F1 vs.
culture-agnostic
metrics

Maximizing F1,
compression
factor (C%)
and
normalized
anti-entropy
(~S) in the
space of hyper-
parameters

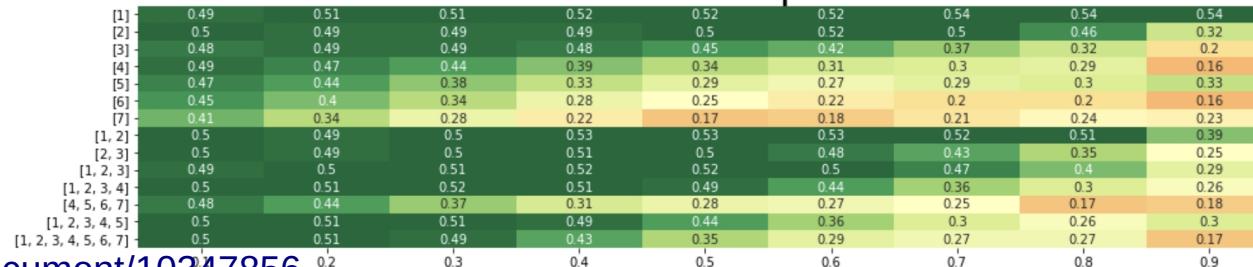
F1:Brown dvf- & dvf+ filter=0 parameters=10967135



C%:Brown dvf- & dvf+ filter=0 parameters=10967135



~S:Brown dvf- & dvf+ filter=0 parameters=10967135

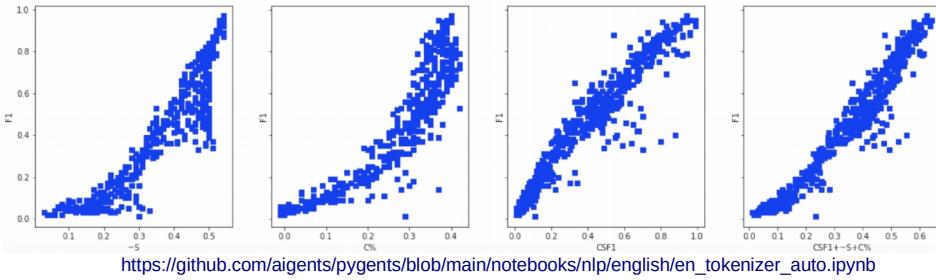


<https://ieeexplore.ieee.org/document/10347856>

https://link.springer.com/chapter/10.1007/978-3-031-44865-2_1

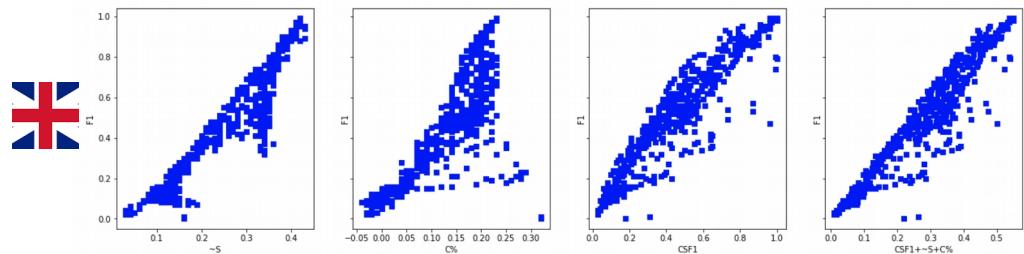
Case 5: Unsupervised Text Segmentation (Different corpora)

English, Train: Brown, Test: Brown 1000

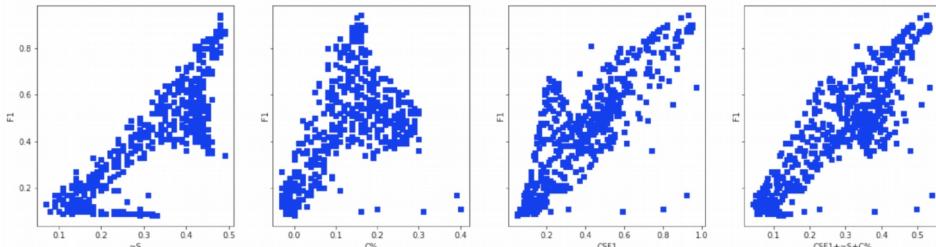


https://github.com/agents/pygents/blob/main/notebooks/nlp/english/en_tokenizer_auto.ipynb

English, Train: Brown, Test: MagicData 100

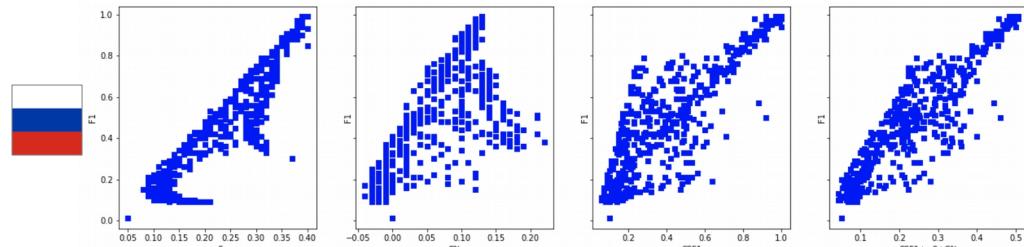


Russian, Train: RusAge, Test: RusAge 1000

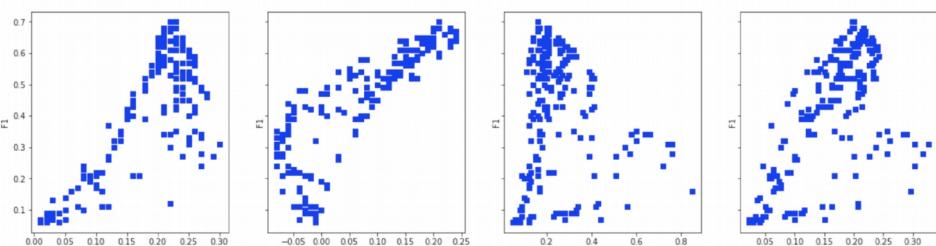


https://github.com/agents/pygents/blob/main/notebooks/nlp/russian/ru_tokenizer_auto.ipynb

Russian, Train: Brown, Test: MagicData 100

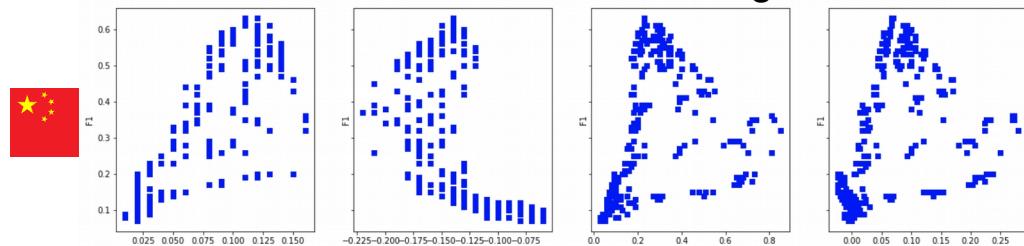


Chinese, Train: CLUE News, Test: CLUE News 1000



https://github.com/agents/pygents/blob/main/notebooks/nlp/chinese/zh_tokenizer_auto.ipynb

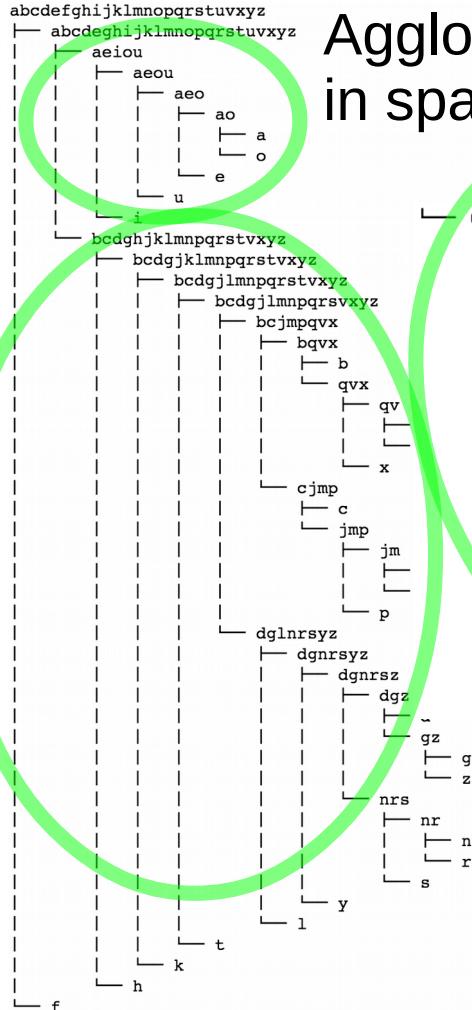
Chinese, Train: Brown, Test: MagicData 100



https://github.com/agents/pygents/blob/main/notebooks/nlp/tokenization/brown/tokenization_brown_en_ru_zh.ipynb

Case 6: Unsupervised Character Category Learning

Agglomerative Clustering in space of Transitions

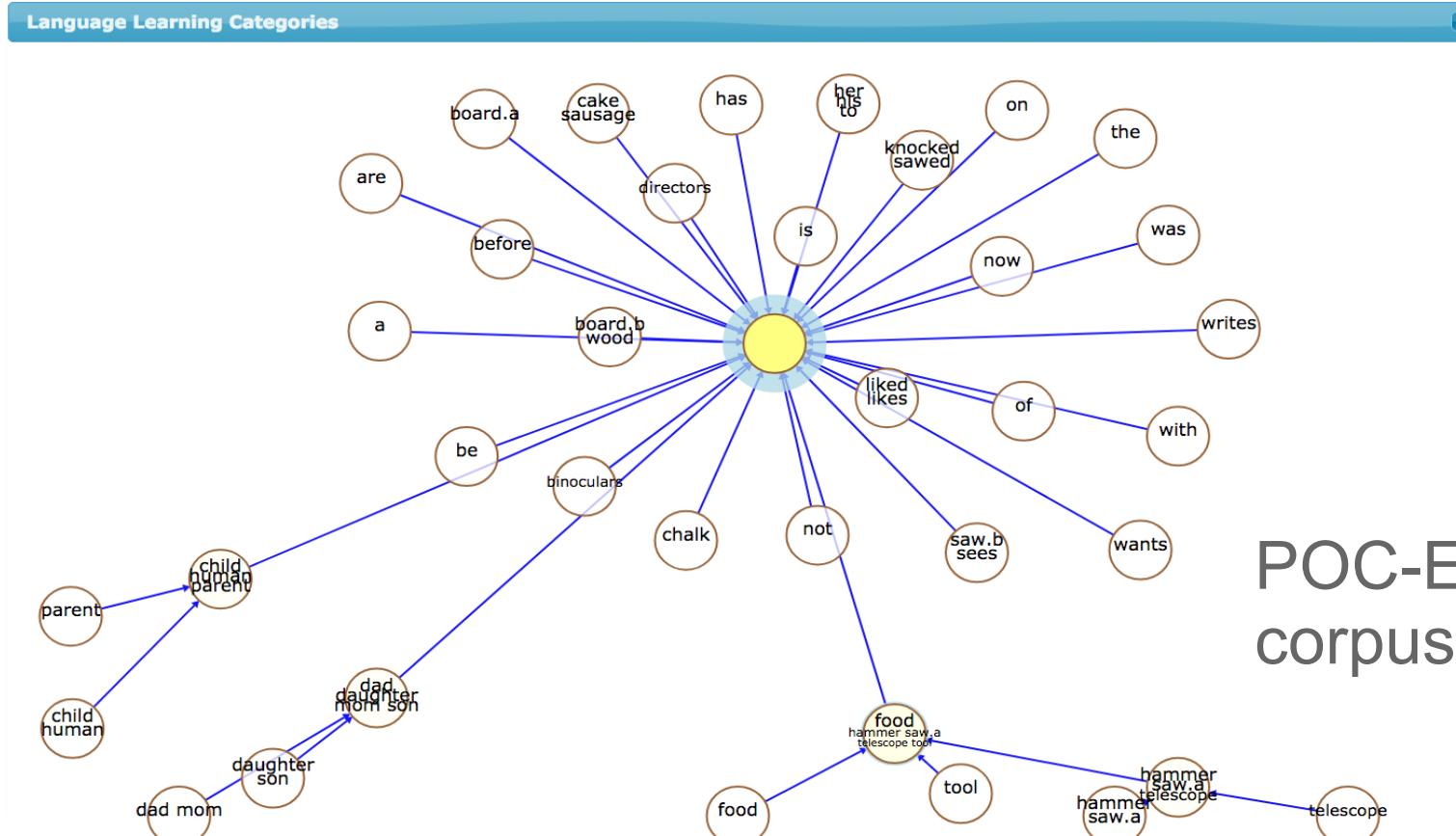


RusAge Test/Small,
Cosine Similarity

Copyright © 2024 Anton Kolonin, Agents®

<https://aclanthology.org/2022.emnlp-main.239.pdf>

Case 7: Unsupervised Ontology Learning on Parses



SingularityNET

<https://www.springerprofessional.de/unsupervised-language-learning-in-opencog/15995030>

<https://www.springerprofessional.de/en/programmatic-link-grammar-induction-for-unsupervised-language-le/17020348>

Thank You for Attention!



Anton Kolonin

akolonin@aigents.com

Messengers: akolonin



<https://agirussia.org>

