Interpretable Natural Language Processing

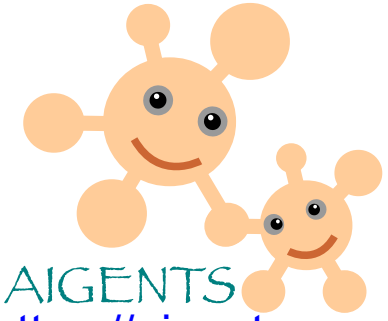
Why and How?

Anton Kolonin

akolonin@aigents.com

Messengers: akolonin

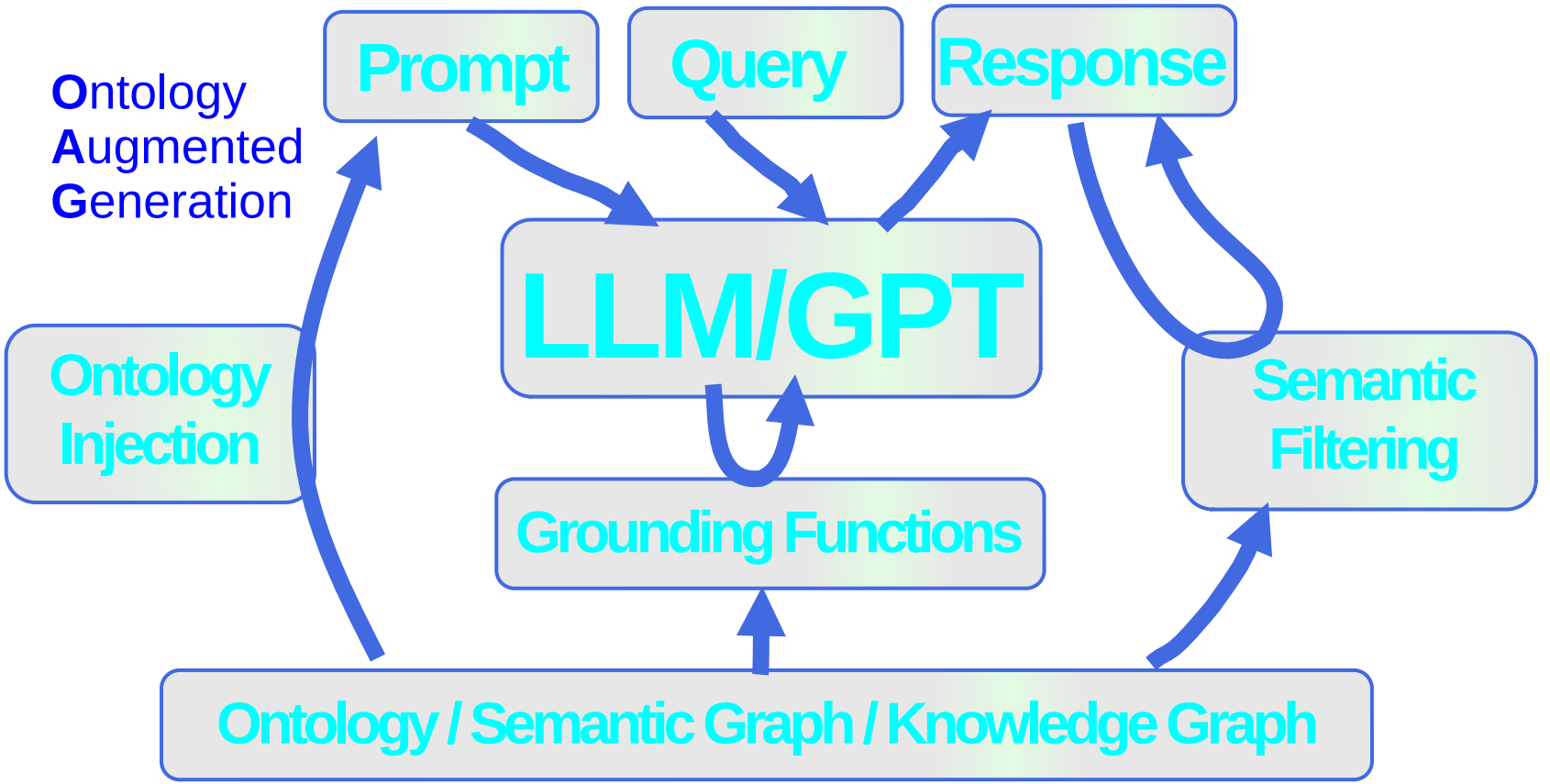
https://agirussia.org



https://aigents.com

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

Fighting Hallucinations with OAG/RAG at Palantir

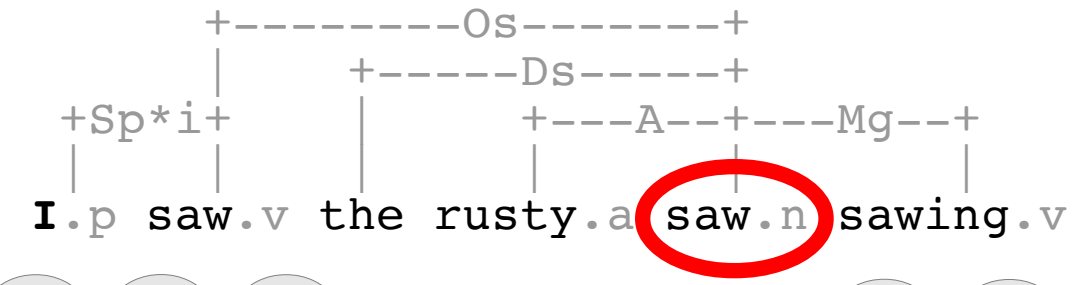


https://blog.palantir.com/reducing-hallucinations-with-the-ontology-in-palantir-aip-288552477383

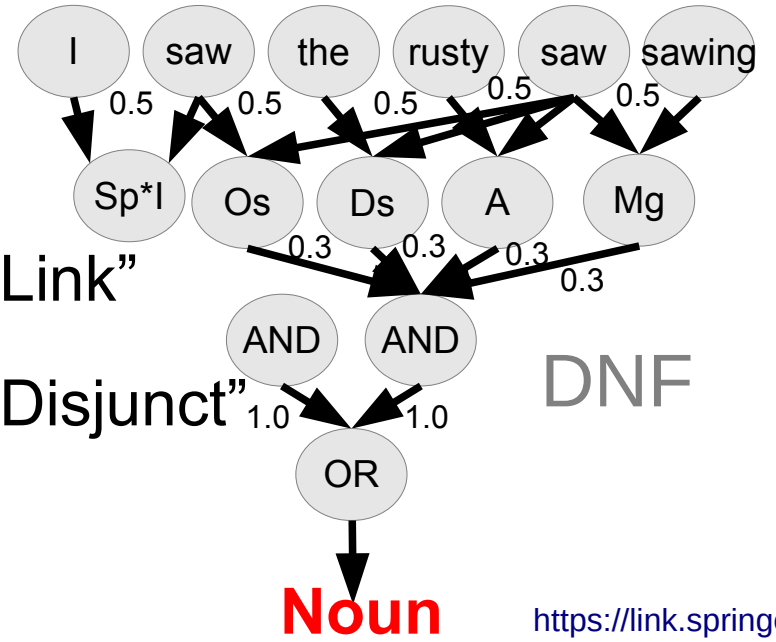
Copyright © 2024 Anton Kolonin, Aigents®

Bridging the Symbolic-Subsymbolic gap in NLP

Formal Link Grammar



Deep Language Model



"Link"

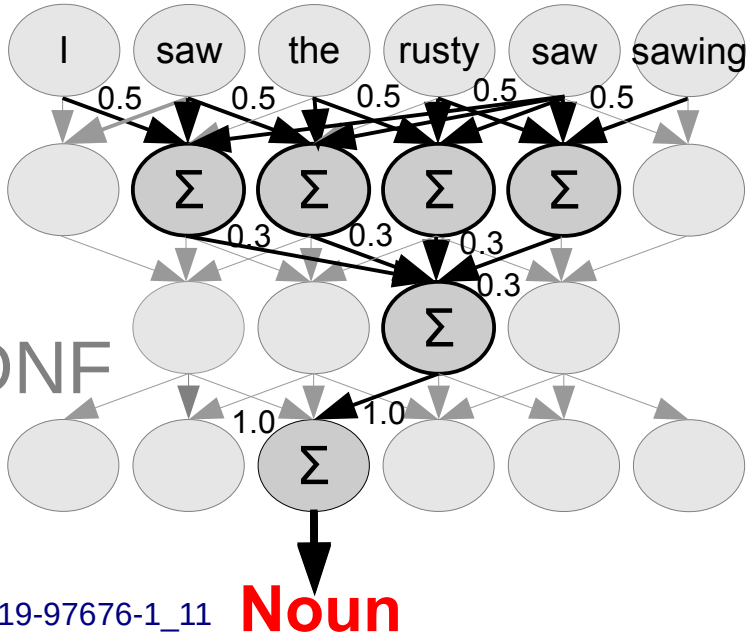
"Disjunct"

DNF

Explain

Transfer

Soft-DNF

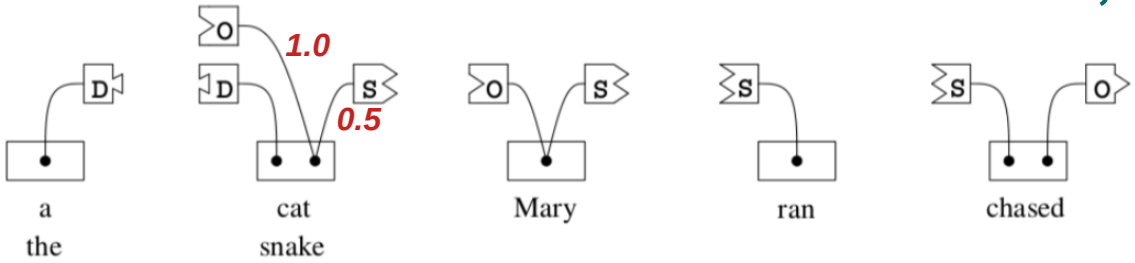


https://link.springer.com/chapter/10.1007/978-3-319-97676-1_11
https://doi.org/10.1007/978-3-030-27005-6_11
https://github.com/singnet/language-learning/

SingularityNET

Copyright © 2024 Anton Kolonin, Aigents®

Link Grammar – Connectors/Costs, Disjuncts

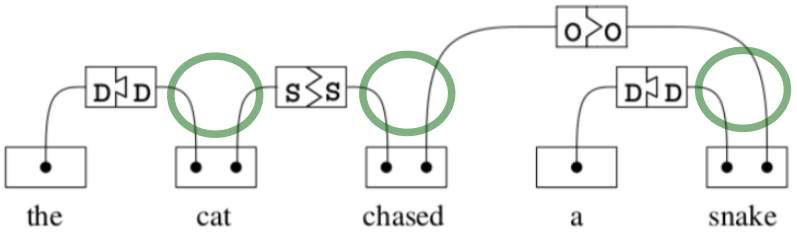


Connectors & Costs

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:



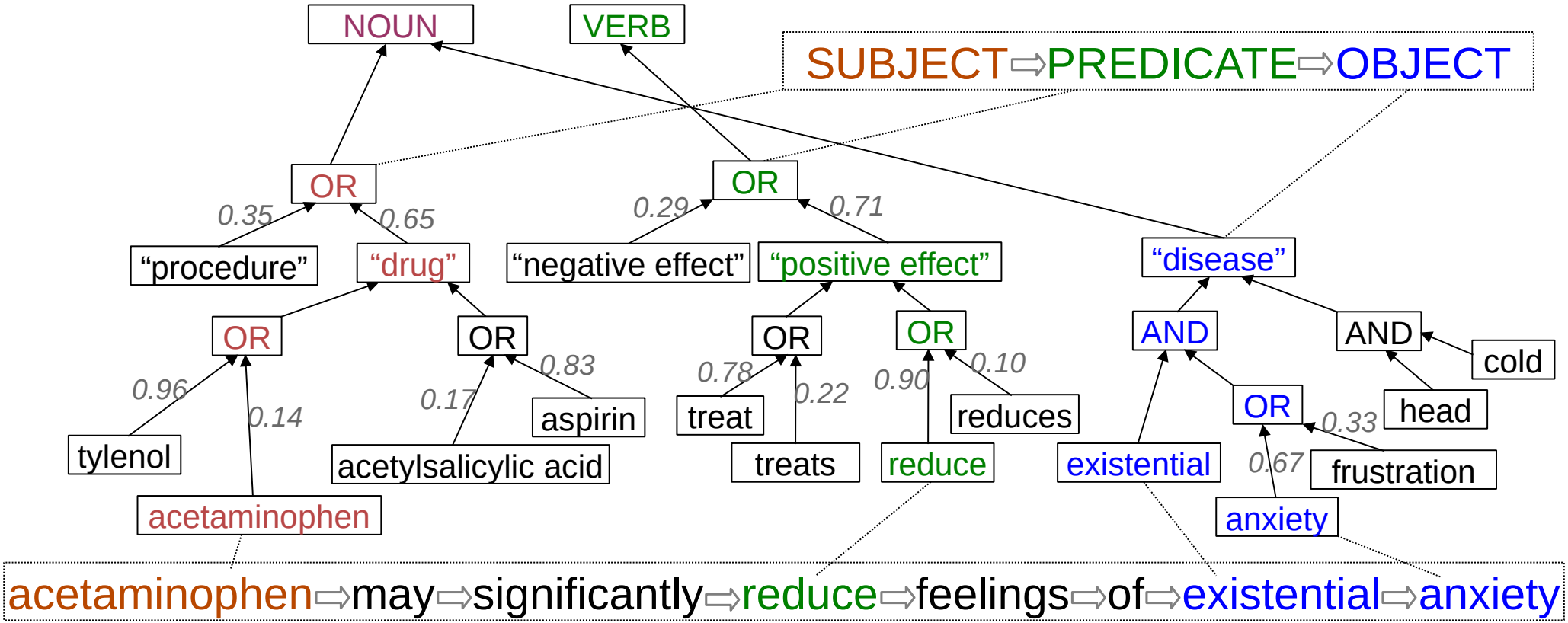
Disjuncts of Connectors

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

SingularityNET

Copyright © 2024 Anton Kolonin, Aigents®

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

Copyright © 2024 Anton Kolonin, Aigents®

Generalized Text Mining with Aigents® "Deep Patterns"

Classification

Category: "Healthcare"

tylenol
acetaminophen
placebo

Case/Relationship Extraction

IS

Entity (Case): "Treatment: Healing anxiety with Tylenol"

significantly
reduce
feelings
study

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

Property Attribution Entity Extraction

HAS

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.

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

Copyright © 2024 Anton Kolonin, Aigents®

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

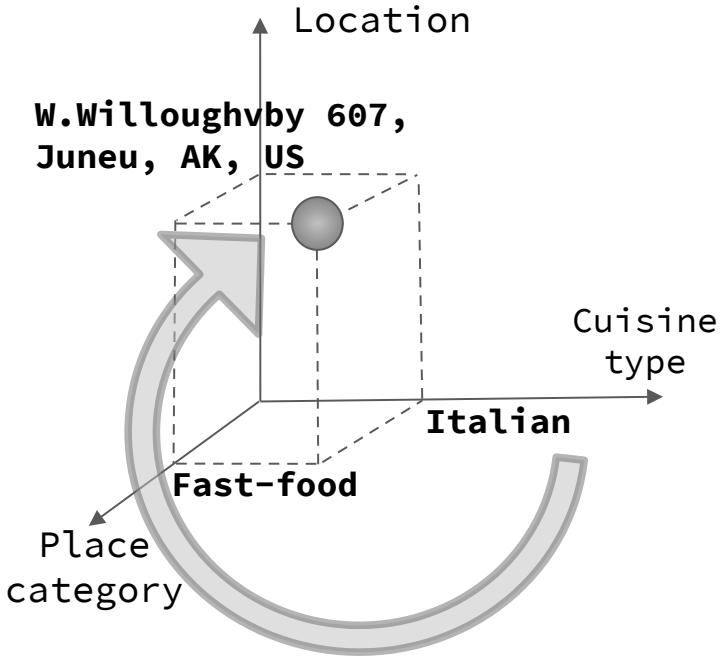
Example:

{[$description catheter] [$coating coating] [$inner-diameter
{diameter inner-diameter}] [$tip tip] [$pattern pattern]}
X
“Convey Guiding Catheter. Unique hydrophilic coating.
Small atraumatic soft tip. Ultra-thin 1 × 2 flat wire braid pattern”
=
{ **coating** : "hydrophillic", **description** : "convey guiding",
pattern : "ultra-thin 1 × 2 flat wire braid", **tip** : "soft" }

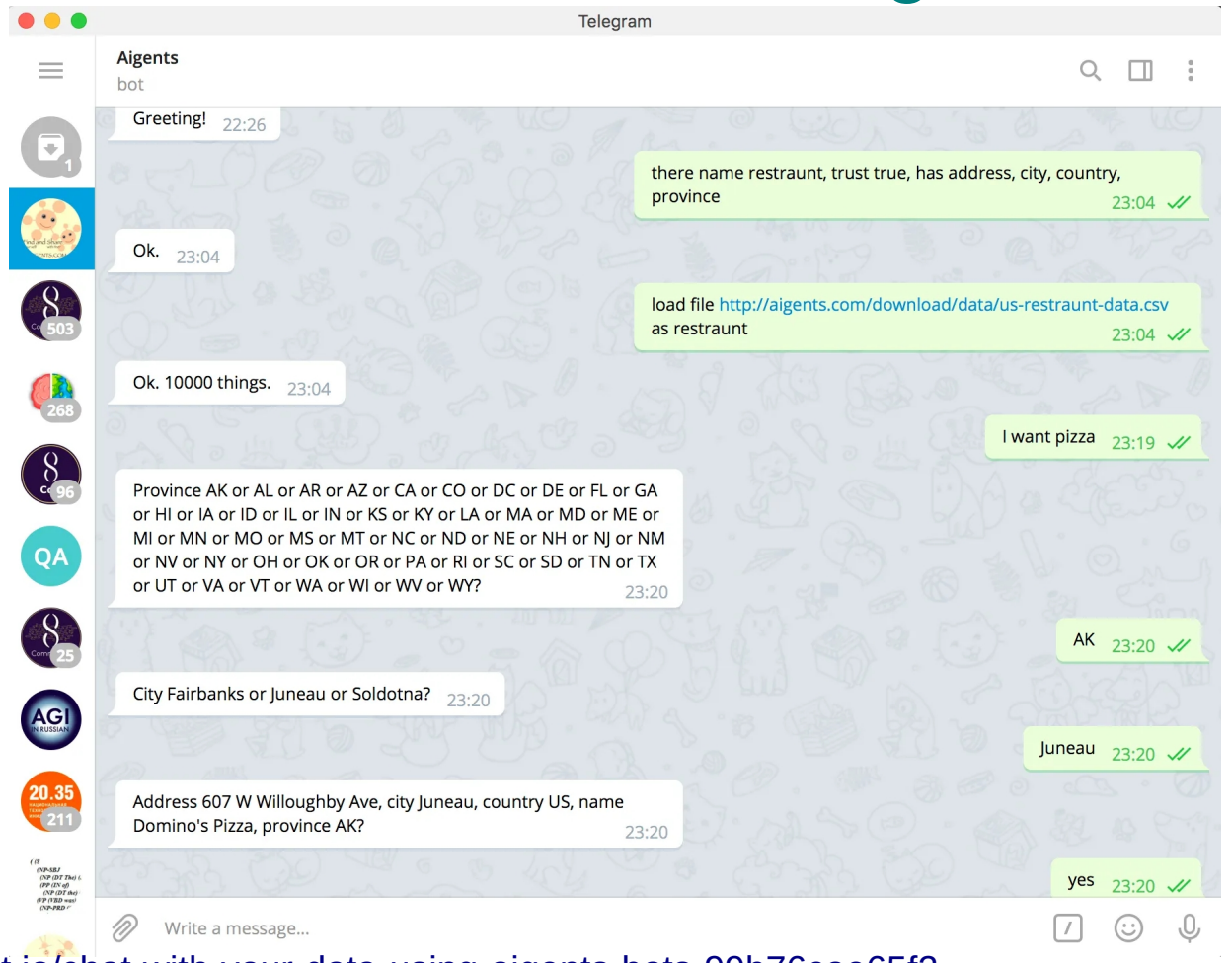
https://ieeexplore.ieee.org/document/7361868
https://github.com/aigents/aigents-java

Copyright © 2024 Anton Kolonin, Aigents®

Case 8: Data-driven Conversational Intelligence



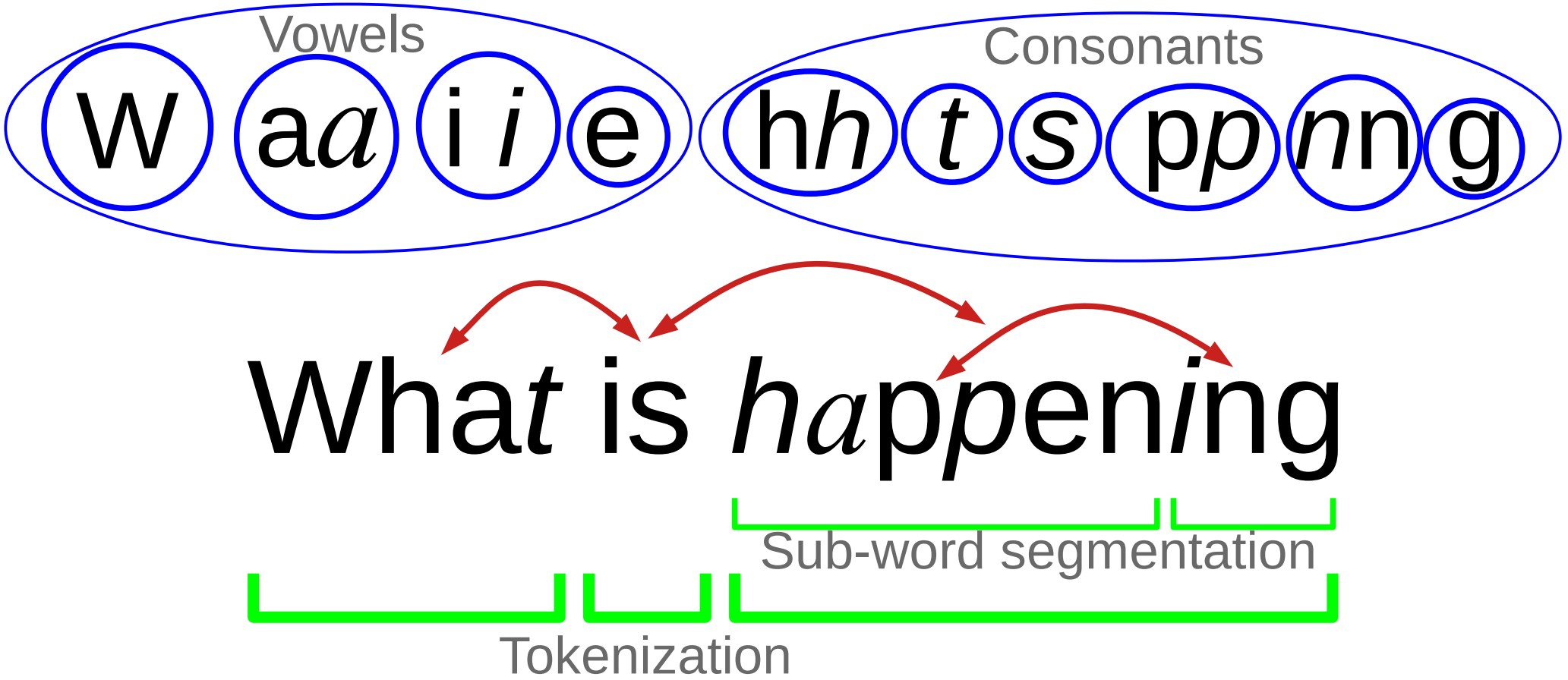
Dynamic data-driven dialog flow minimizing uncertainty (Can be used for OAG/TAG)



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

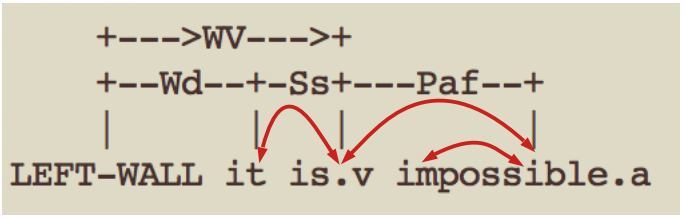
Copyright © 2024 Anton Kolonin, Aigents®

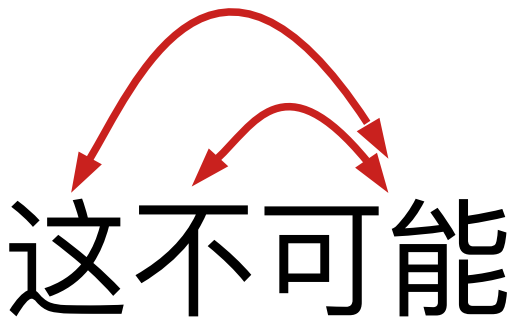
Learning Lexicon, Punctuation, Morphology and Grammar
Clustering, Segmentation and Parsing



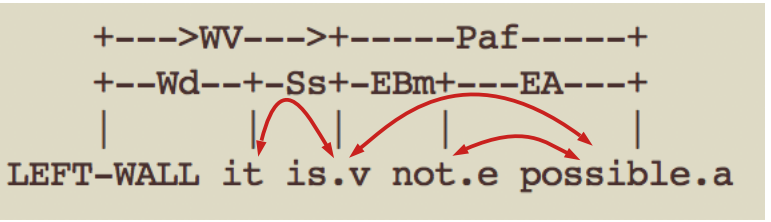
Copyright © 2024 Anton Kolonin, Aigents®

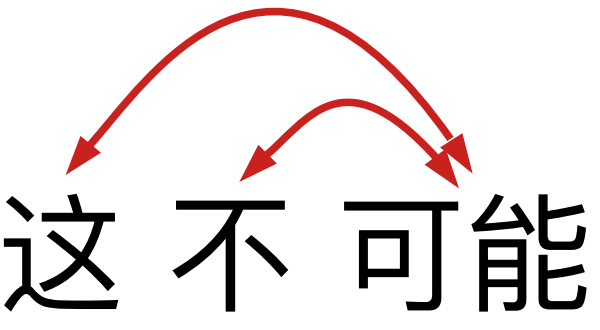
Blurring boundaries between grammatical parsing and morphological parsing causes tokenization ambiguity





(Google)

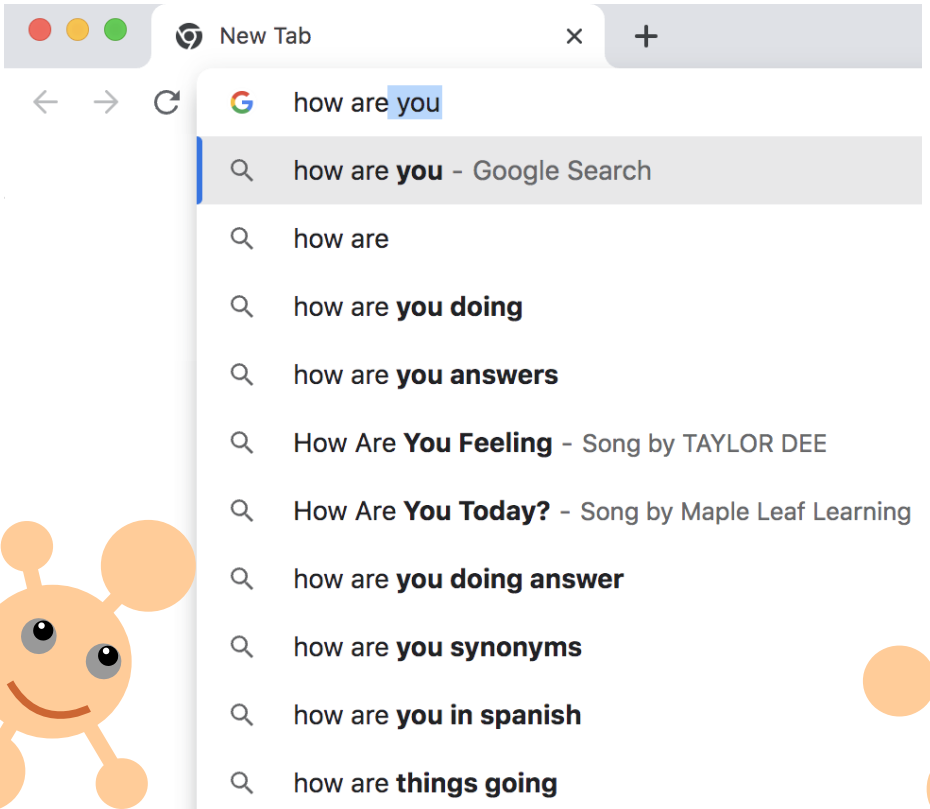


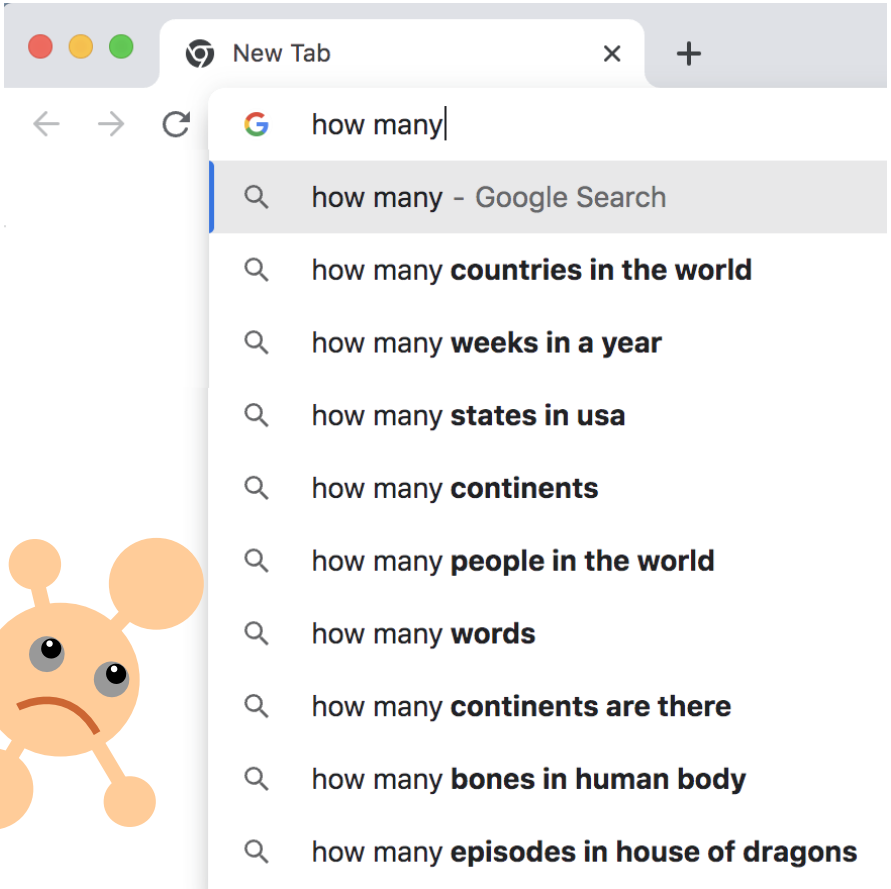


(Jieba)

Copyright © 2024 Anton Kolonin, Aigents®

Minimizing Uncertainty

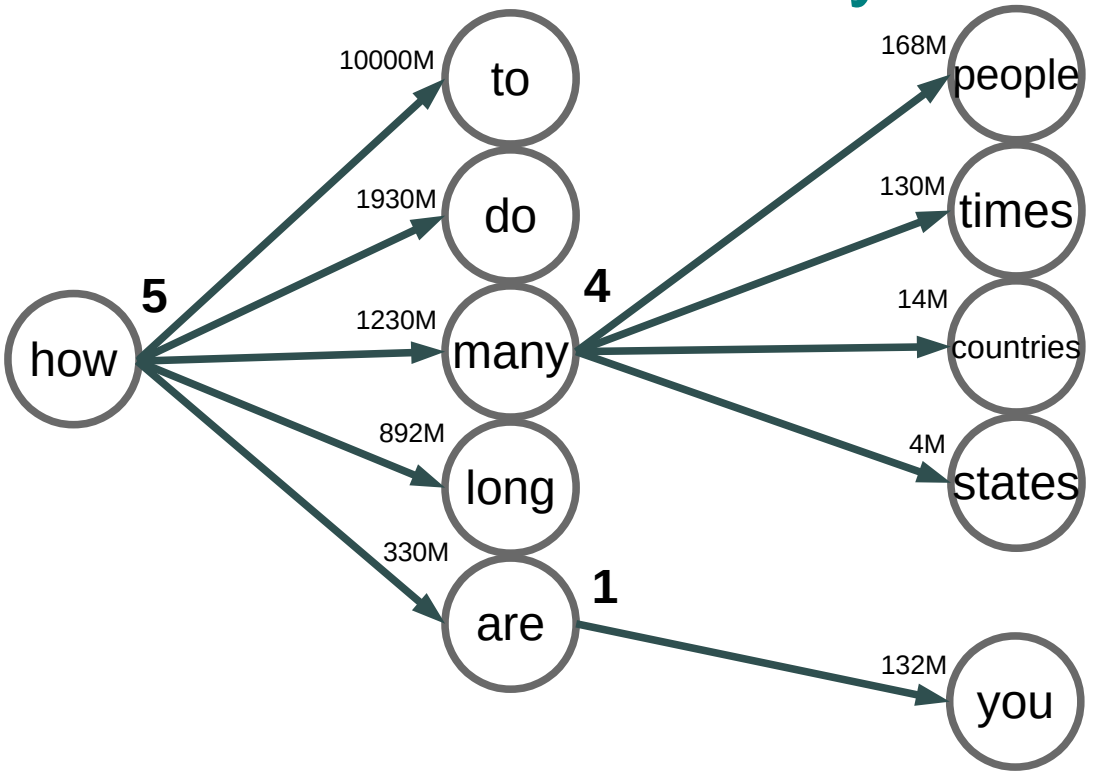




https://aclanthology.org/2022.emnlp-main.239/
https://arxiv.org/pdf/2303.02427.pdf

Copyright © 2024 Anton Kolonin, Aigents®

Unsupervised Learning for Text Segmentation based on Probability and Uncertainty Measures



Metrics/Indicators:

Mutual Information[1]
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

Copyright © 2024 Anton Kolonin, Aigents®

Case 5: Unsupervised Text Segmentation (Probability)

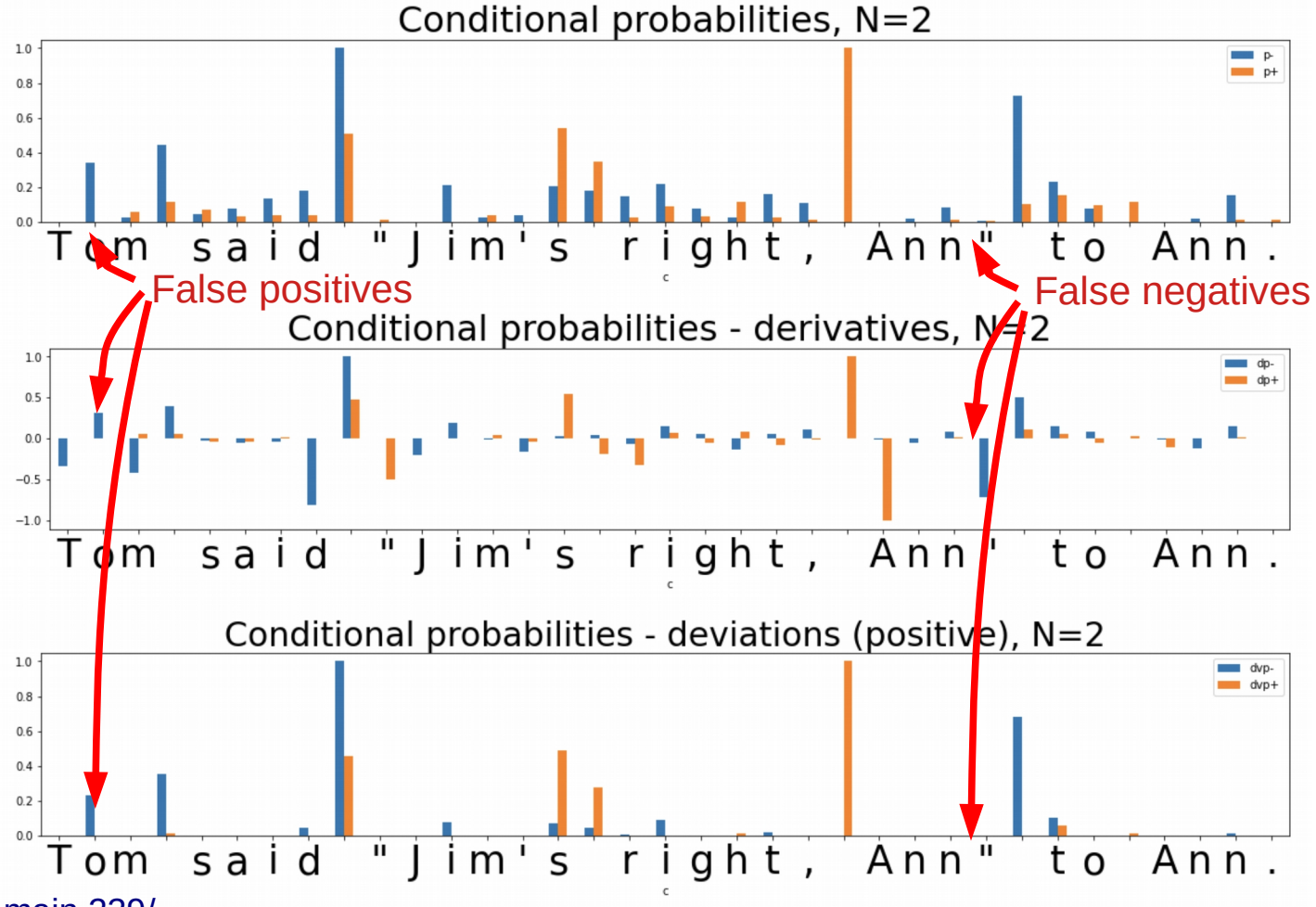
Metrics/Indicators:

Ngram (Character) Conditional Probability

(of Transition)

$P(Ngram_{n+1})/P(Ngram_n)$

P(“m ”)/P(“m”)



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

Copyright © 2024 Anton Kolonin, Aigents®

Case 5: Unsupervised Text Segmentation (Uncertainty)

Metrics/ Indicators:

Transition
Freedom
Deviation

(varying “N” of N-gram)

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

Copyright © 2024 Anton Kolonin, Aigents®

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

Copyright © 2024 Anton Kolonin, Aigents®

Case 5: Unsupervised Text Segmentation (Hyper-parameters)

English

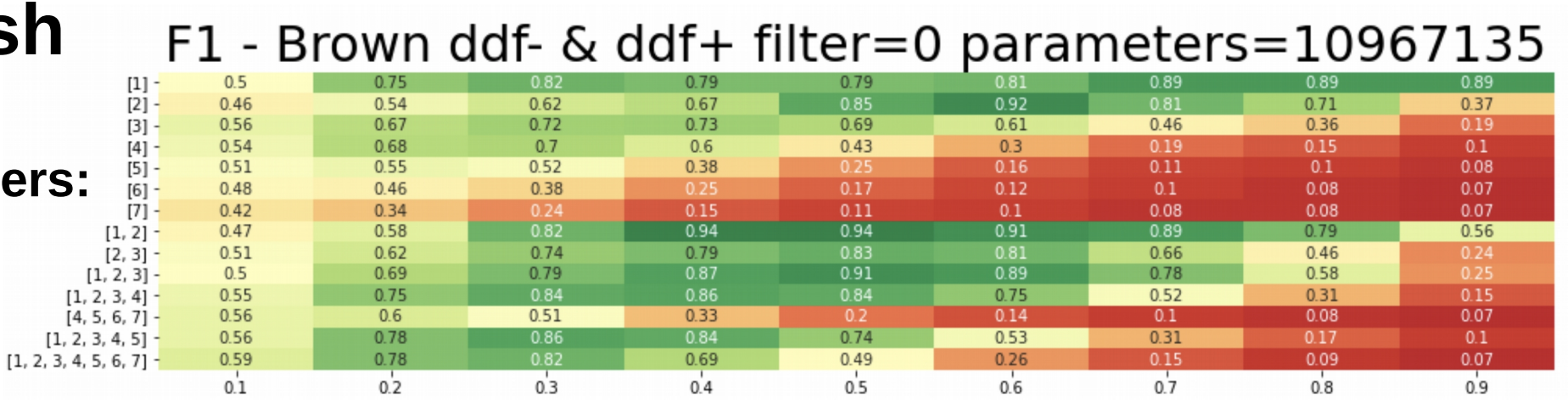
Hyper-Parameters:

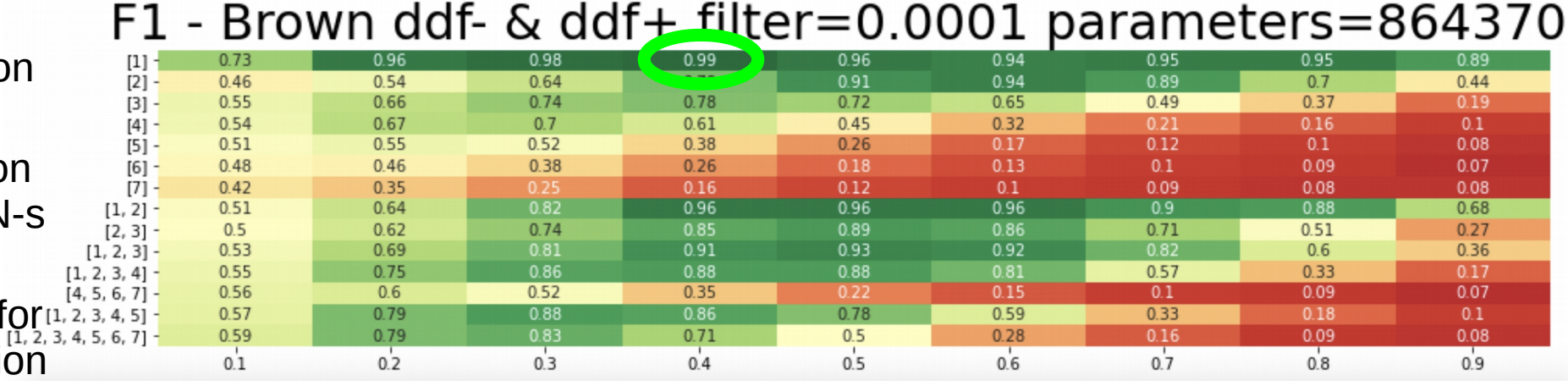
Metric: Transition Freedom

Threshold for model compression

Combination of Ngram N-s

Threshold for segmentation





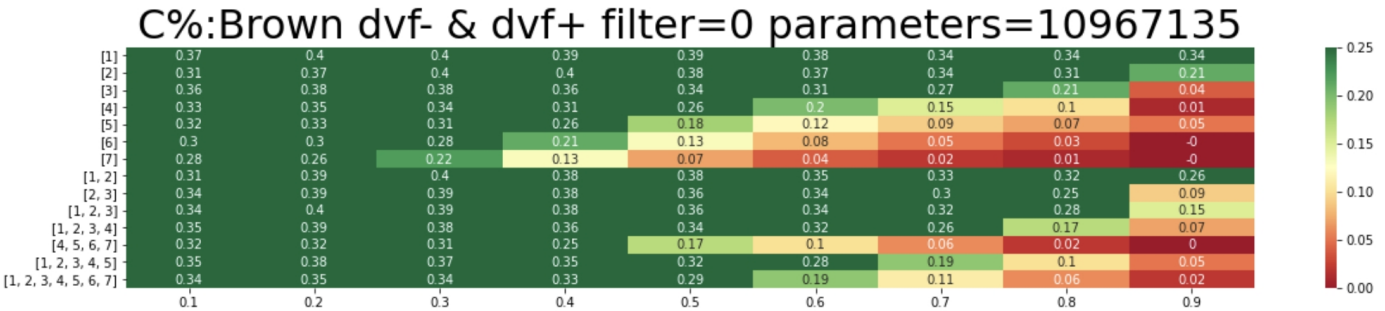
https://aclanthology.org/2022.emnlp-main.239/

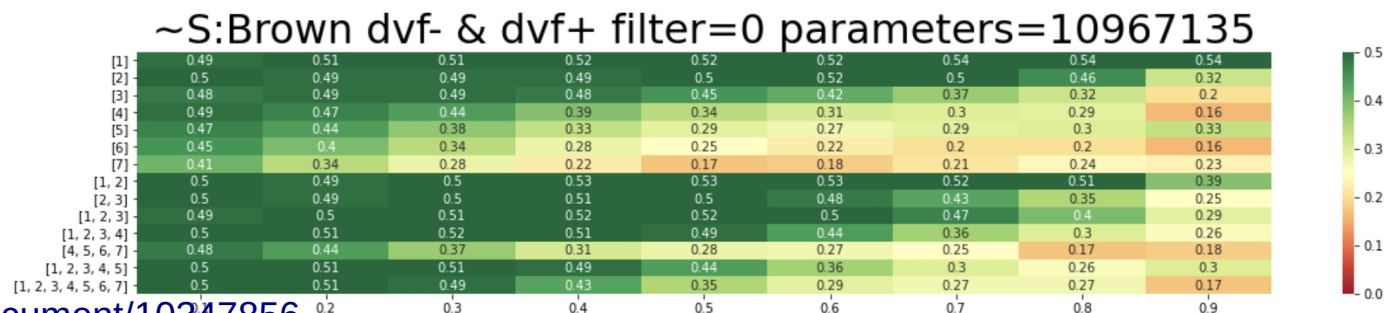
Copyright © 2024 Anton Kolonin, Aigents®

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



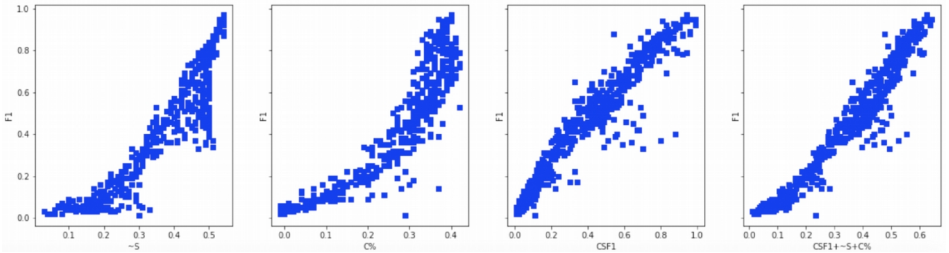


https://ieeexplore.ieee.org/document/10347856
https://link.springer.com/chapter/10.1007/978-3-031-44865-2_1

Copyright © 2024 Anton Kolonin, Aigents®

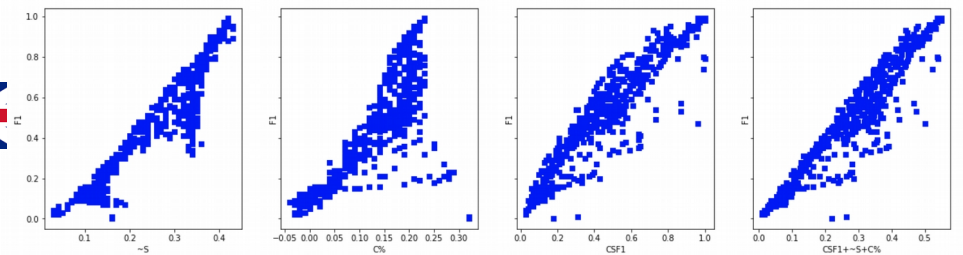
Case 5: Unsupervised Text Segmentation (Different corpora)

English, Train: Brown, Test: Brown 1000

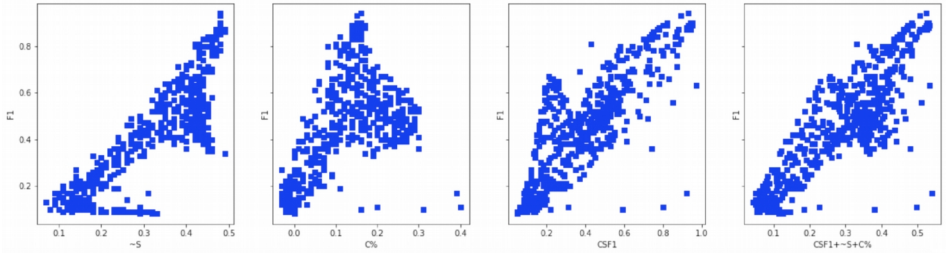


https://github.com/aigents/pygents/blob/main/notebooks/nlp/english/en_tokenizer_auto.ipynb

English, Train: Brown, Test: MagicData 100

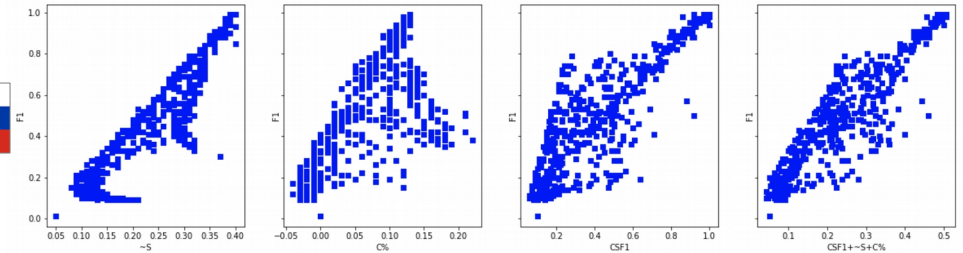


Russian, Train: RusAge, Test: RusAge 1000

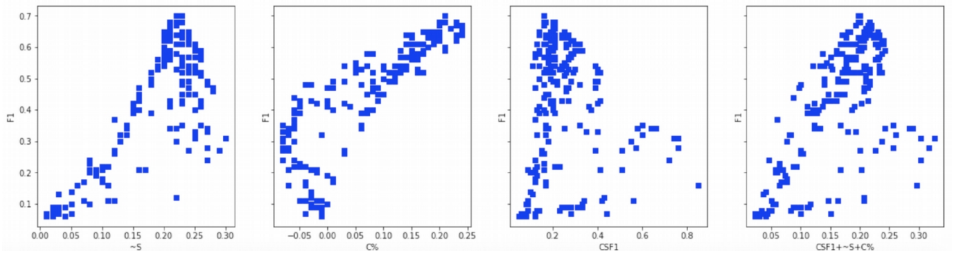


https://github.com/aigents/pygents/blob/main/notebooks/nlp/russian/ru_tokenizer_auto.ipynb

Russian, Train: Brown, Test: MagicData 100

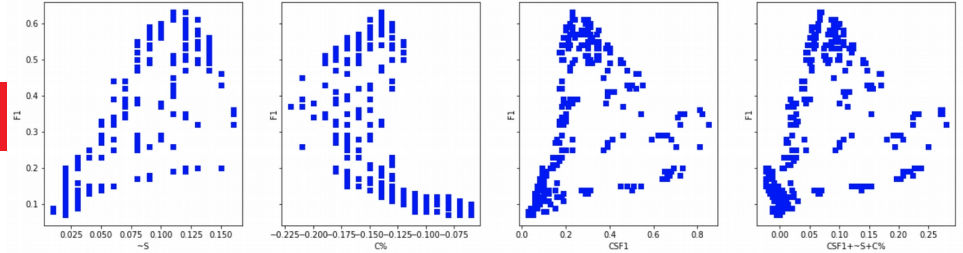


Chinese, Train: CLUE News, Test: CLUE News 1000



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

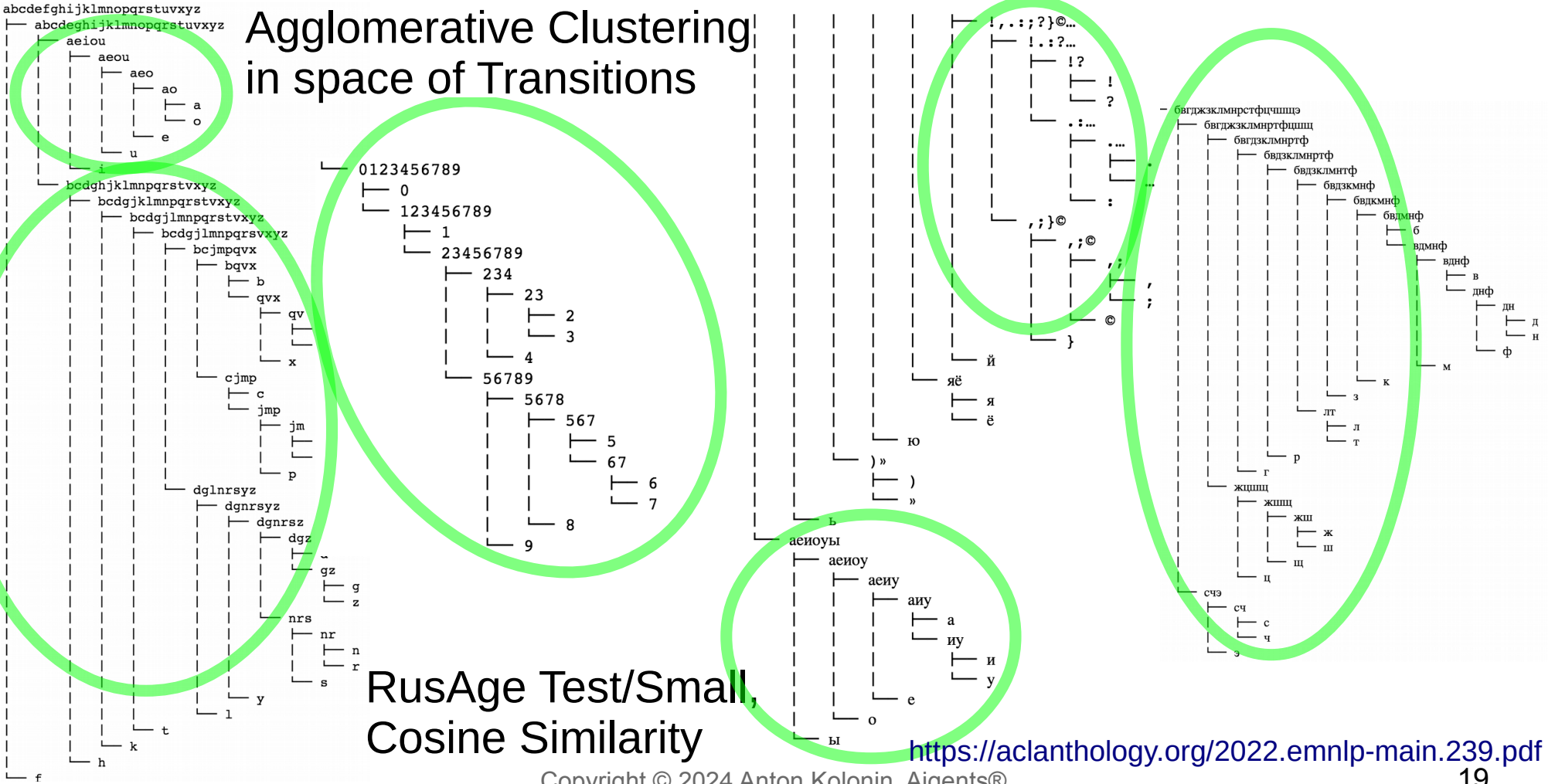
Chinese, Train: Brown, Test: MagicData 100



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

Copyright © 2024 Anton Kolonin, Aigents®

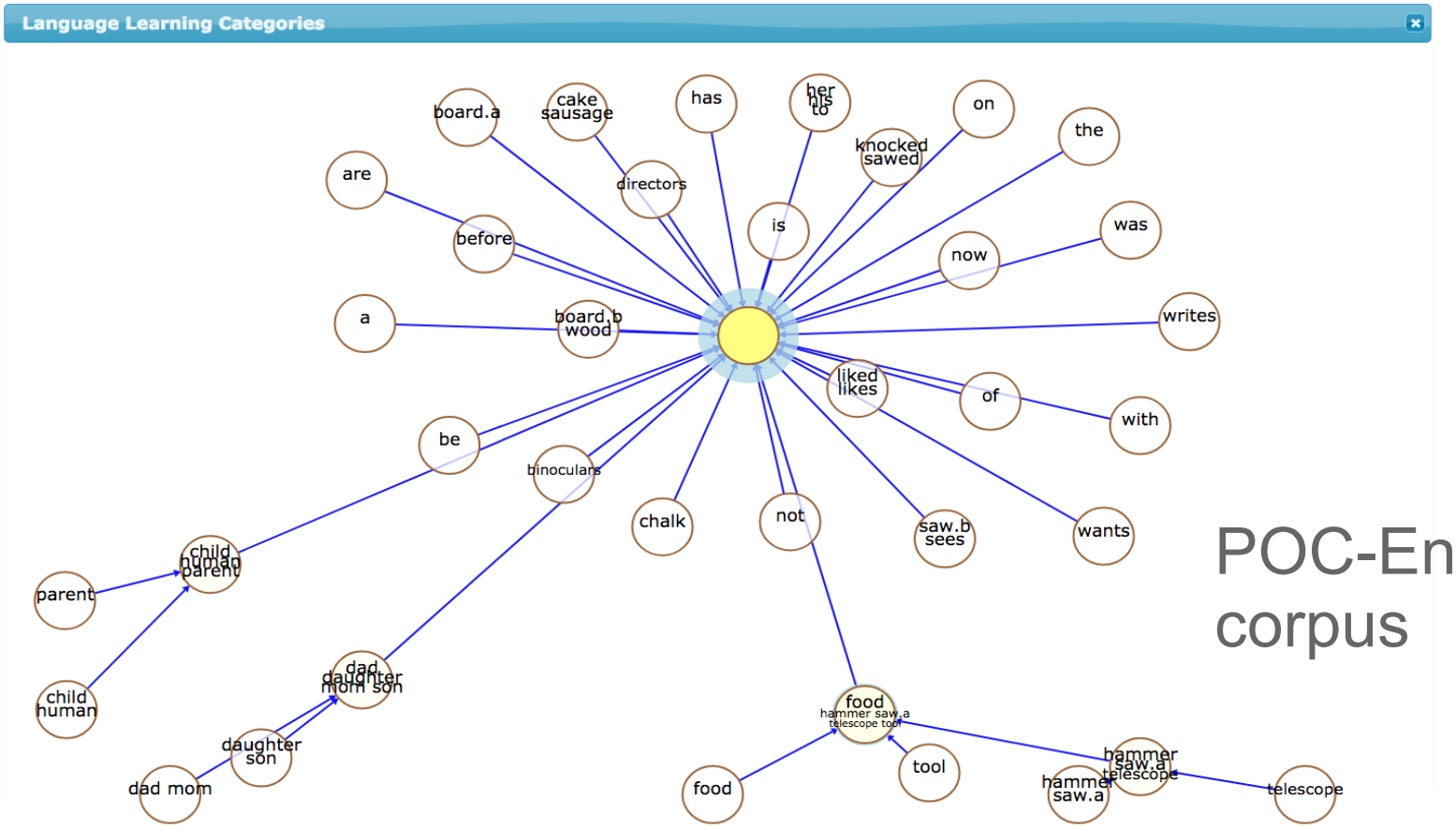
Case 6: Unsupervised Character Category Learning



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

Copyright © 2024 Anton Kolonin, Aigents®

Case 7: Unsupervised Ontology Learning on Parses



POC-English corpus

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

Copyright © 2024 Anton Kolonin, Aigents®

Thank You for Attention!

Anton Kolonin

akolonin@aigents.com

Messengers: akolonin

https://agirussia.org

AIGENTS

https://aigents.com

Novosibirsk State University

*THE REAL SCIENCE

SingularityNET

Copyright © 2024 Anton Kolonin, Aigents®