

Interpretable Natural Language Processing

Fundamental studies for applied results



<https://agirussia.org>

Anton Kolonin
akolonin@aigents.com
Messengers: akolonin



The Plan

A bit of theory

Few cases

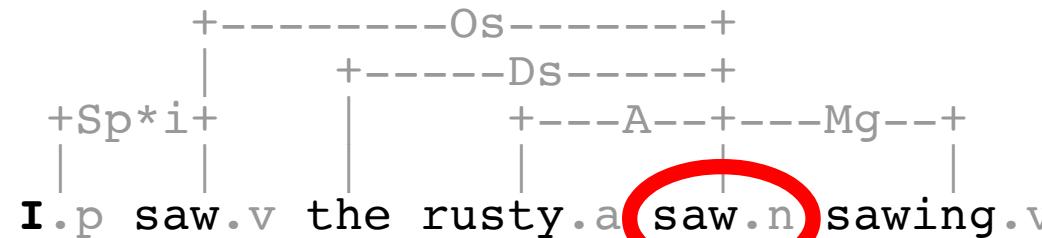
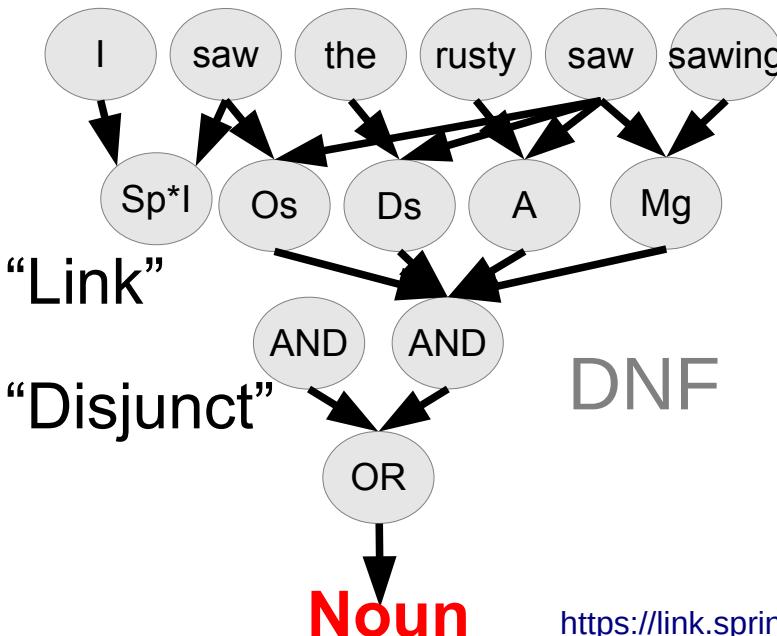
A bit more theory

A few more cases

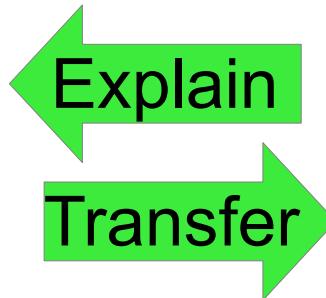
A bit of theory:
Fuzzy/probabilistic/soft-logic symbolic “models”

Bridging the Symbolic-Subsymbolic gap in NLP

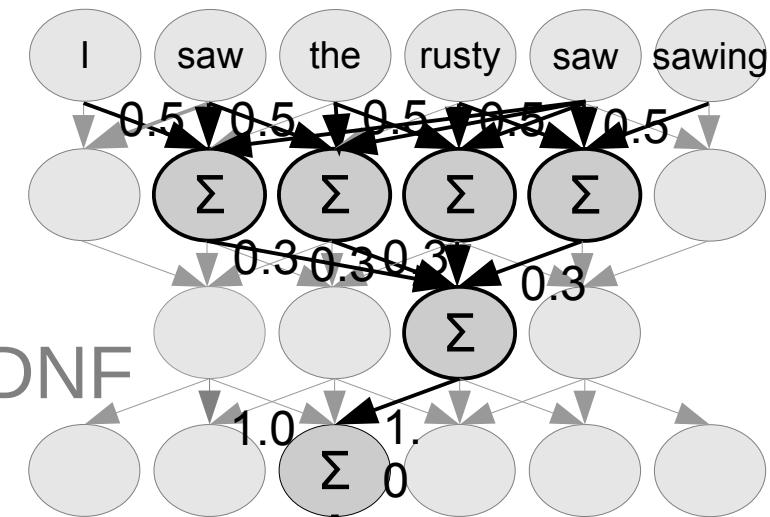
Formal
Link Grammar



Deep
Language Model



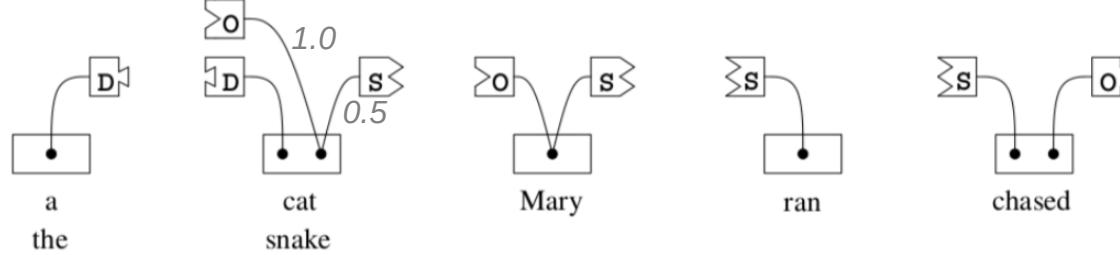
Soft-DNF



Noun

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/>

Link Grammar – Connectors/Costs, Disjuncts

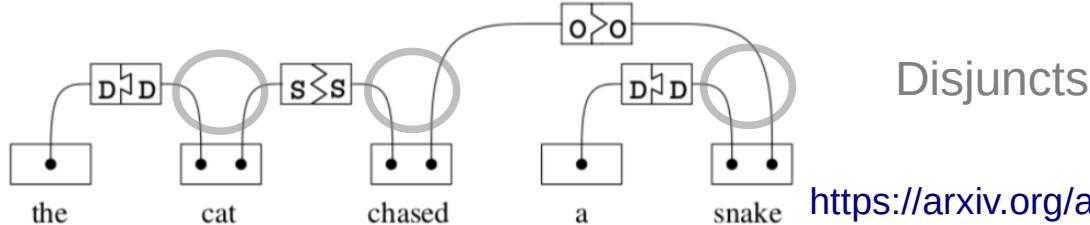


Connectors/Costs

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+;
```

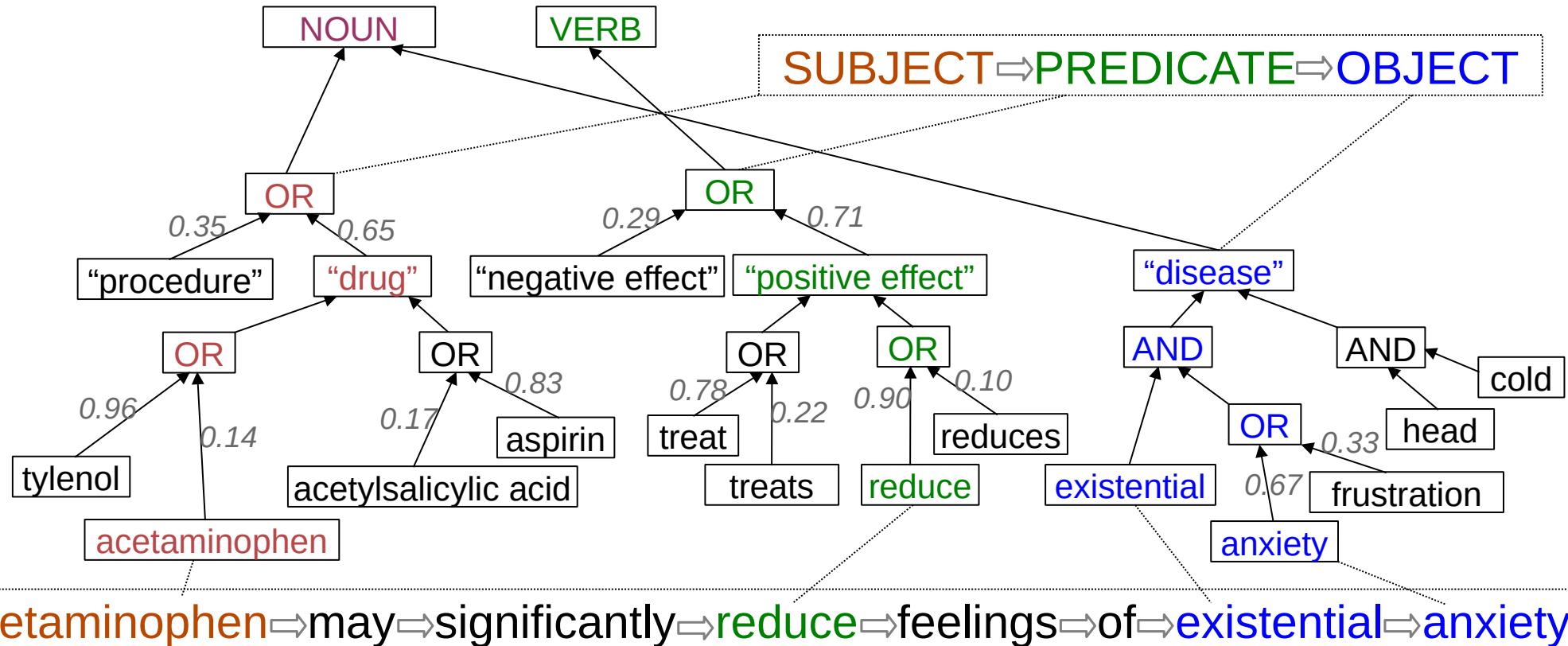
Note that although the symbols ‘‘&’’ and ‘‘or’’ are used to write down disjuncts, these are **not** Boolean operators, and do **not** form a Boolean algebra. They do form a non-symmetric compact closed monoidal algebra. The diagram below illustrates puzzle pieces, assembled to form a parse:



Disjuncts

<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.

Few cases: Using symbolic “models”

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”

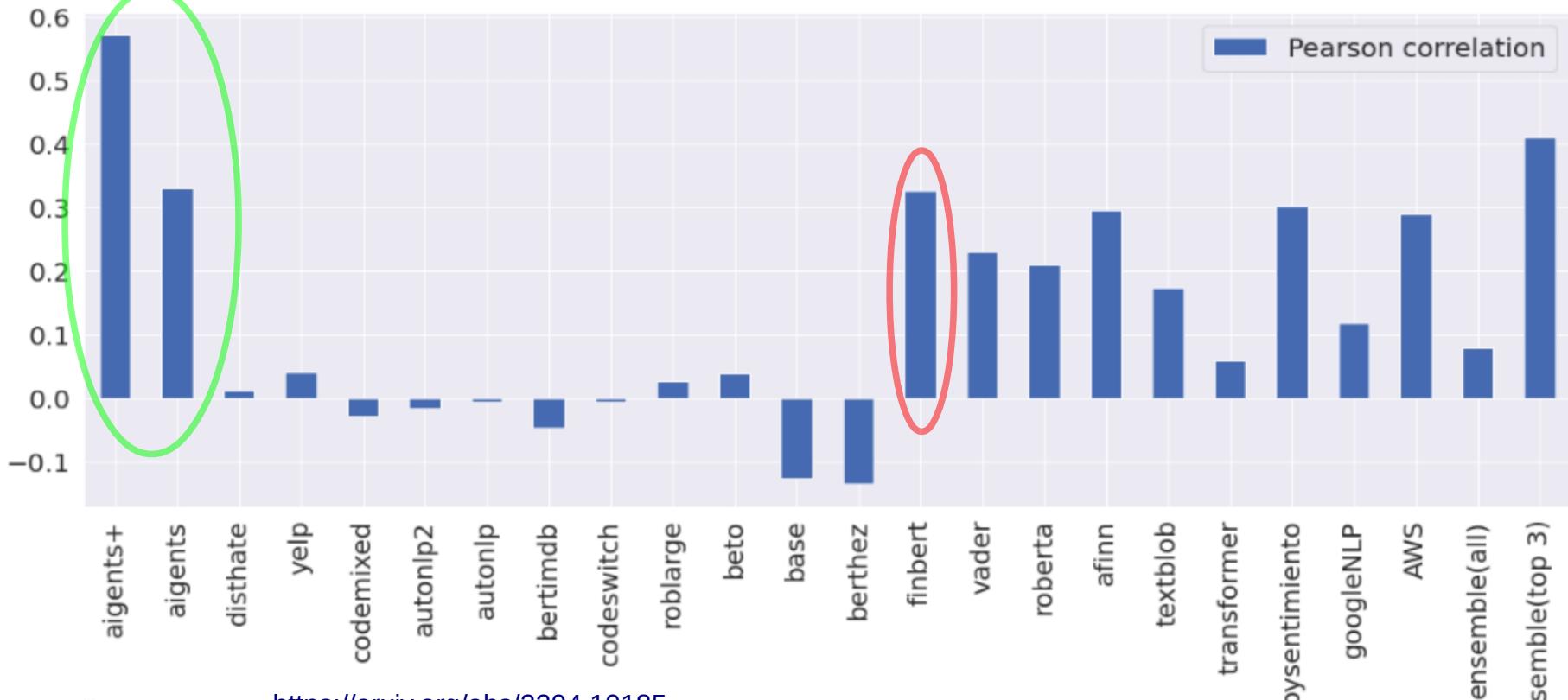
=

```
{ 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 2: Interpretable Sentiment Analysis

Aigents® “interpretable” model vs. Bert fine-tuned on financial data
Average correlation across all models



Case 2.1: Aigents® News Feed Sentiment Analysis

The screenshot shows the Aigents news feed interface. At the top, there's a header with the Aigents logo, user profile 'Madmind Siberian', and various social media and utility icons. Below the header is a search bar labeled 'Input search text' with a '+' button. The main area displays a list of news items from different sources, each with a timestamp (today, yesterday), a thumbnail image, a brief summary, and a URL. Each news item has a small colored bar indicating its sentiment score, ranging from yellow (positive) to red (negative). The news items include:

- today: putin says he has noted joe biden's sharp anti-Russian rhetoric (red)
- today: the united states of america is set for a pivotal election where joe bide current president donald trump (yellow)
- yesterday: **learn more about blockchain domains:** introduction video 1 see you on thursday (green)
- yesterday: The Russians may know more than we do : cia analyst says putin I doctors amid covid secrecy (red)
- yesterday: alexa-derived chatbot tech allegedly understands the british accent (green)
- yesterday: but not strongly enough to make a difference published: 16 mar 2 bernie sanders lost his last chance to take joe biden down (red)
- yesterday: it must also support in-space assembly manufacturing of hardware be tested and evaluated for future use (green)

At the bottom of the page, there's a copyright notice: 'Copyright 2020 IP Anton Kolonin, Aigents®, Privacy Policy'.

This screenshot shows the Aigents news feed interface, similar to the one on the left but with a different set of news items. The layout is identical, featuring a header with the Aigents logo, user profile 'Madmind Siberian', and social media icons. Below the header is a search bar labeled 'Input new thing name or template' with a '+' button. The main area displays a list of news items with their timestamps, thumbnails, summaries, URLs, and sentiment bars. The news items include:

- biden (red)
- jinpingle (green)
- online social help (yellow)
- personal artificial intelligence (red)
- putin (red)
- sanders (red)
- trump (red)
- {agi hlai [strong ai] [strong artificial intelligence] [artificial general intelligence]} (yellow)
- {chatbot chat-bot [chat bot]} (green)
- {published putin vladimir} (yellow)
- {bitcoin btc btc/usd btcusd} (yellow)
- {dollar usd usd/xau usdxau} (yellow)
- {ethereum eth eth/usd ethusd} (yellow)
- {gold xau xau/usd xauusd} (yellow)

At the bottom of the page, there's a copyright notice: 'Copyright 2020 IP Anton Kolonin, Aigents®, Privacy Policy'.

<https://blog.singularitynet.io/aigents-sentiment-detection-personal-and-social-relevant-news-be989d73b381>

Case 2.2: Aigents® Bot Sentimental Reactions in Chats

AigentsTest
15 members, 2 online

← Aigents Test

Life is wonderful 21:51 ✓

Aigents Bot Admin
Anton Kolonin
Life is wonderful
21:51

Death is not good 21:52 ✓

Aigents Bot Admin
Anton Kolonin
Death is not good
21:52

When I am in the bad mood, it sucks 21:53 ✓

Aigents Bot Admin
Anton Kolonin
When I am in the bad ...
21:53

Write a message... <https://t.me/AigentsBot>

Aigents
bot

← Aigents

5Ogm0fxNSEL7YoRYmvzmxo9U.jpg¹
title: biden to nominate ex-iowa gov
[https://www.reddit...](https://www.reddit.com/r/JoeBiden)

reddit
President-elect Joe Biden • r/JoeBiden
President-elect Joe Biden | We are the United States of America. There is not a single thing we cannot do. Are you with...

BID HAR

17:51

December 11

my telegram report, period 4 21:37 ✓

report.html
40.8 KB
21:37

Forwarded from Anton Kolonin
Deaths sucks 21:51

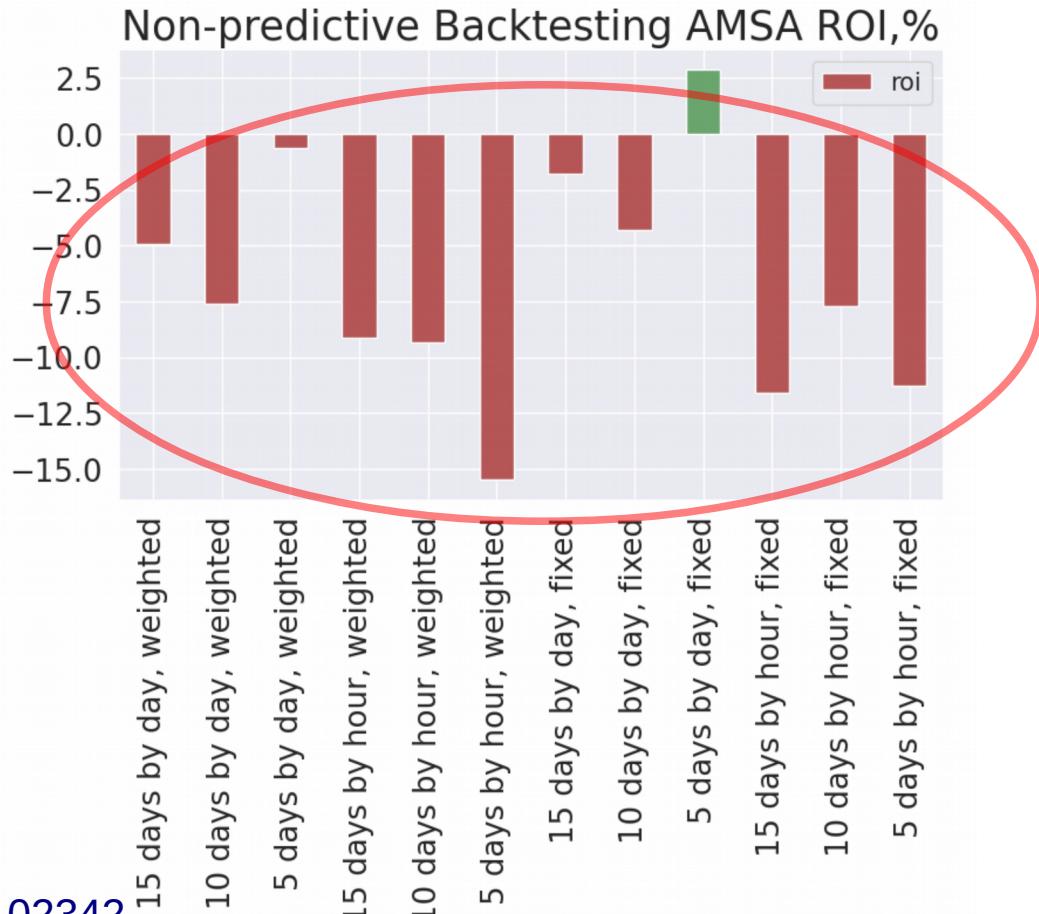
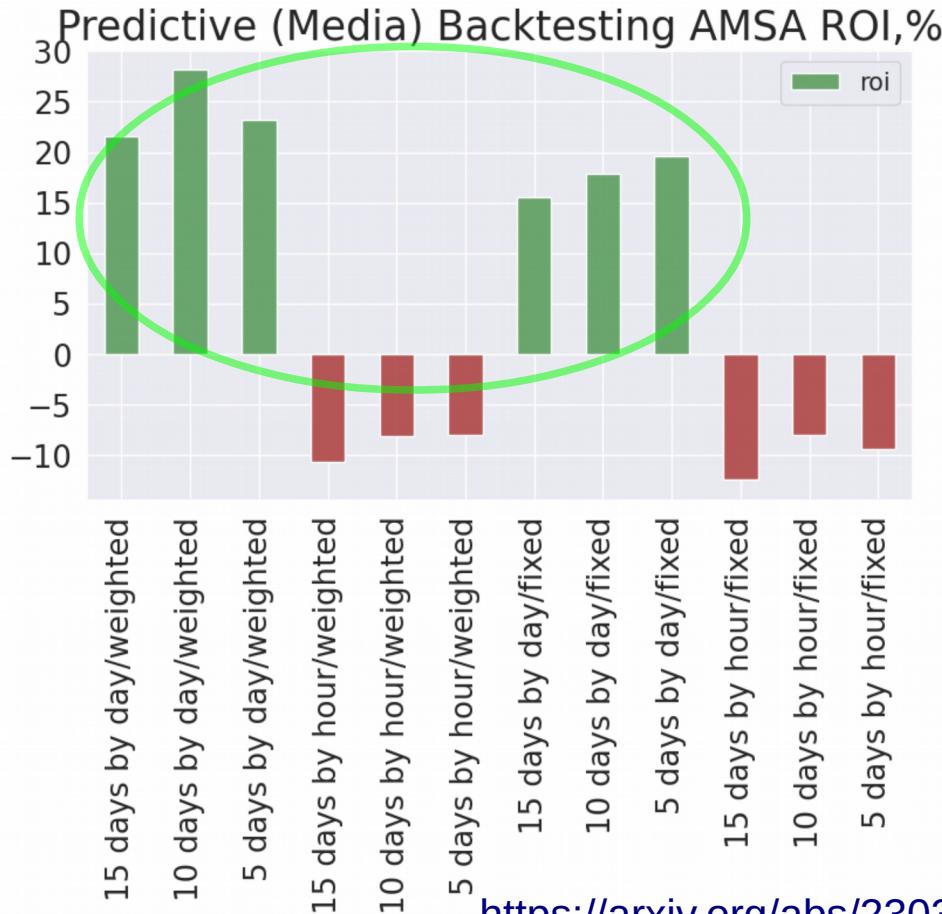
Forwarded from Anton Kolonin
When I am in the bad mood, it sucks 21:53

Write a message... <https://agents.medium.com/aigents-bot-for-telegram-groups-1dba32140047>

Case 3: Cognitive Distortions on Media vs. Market



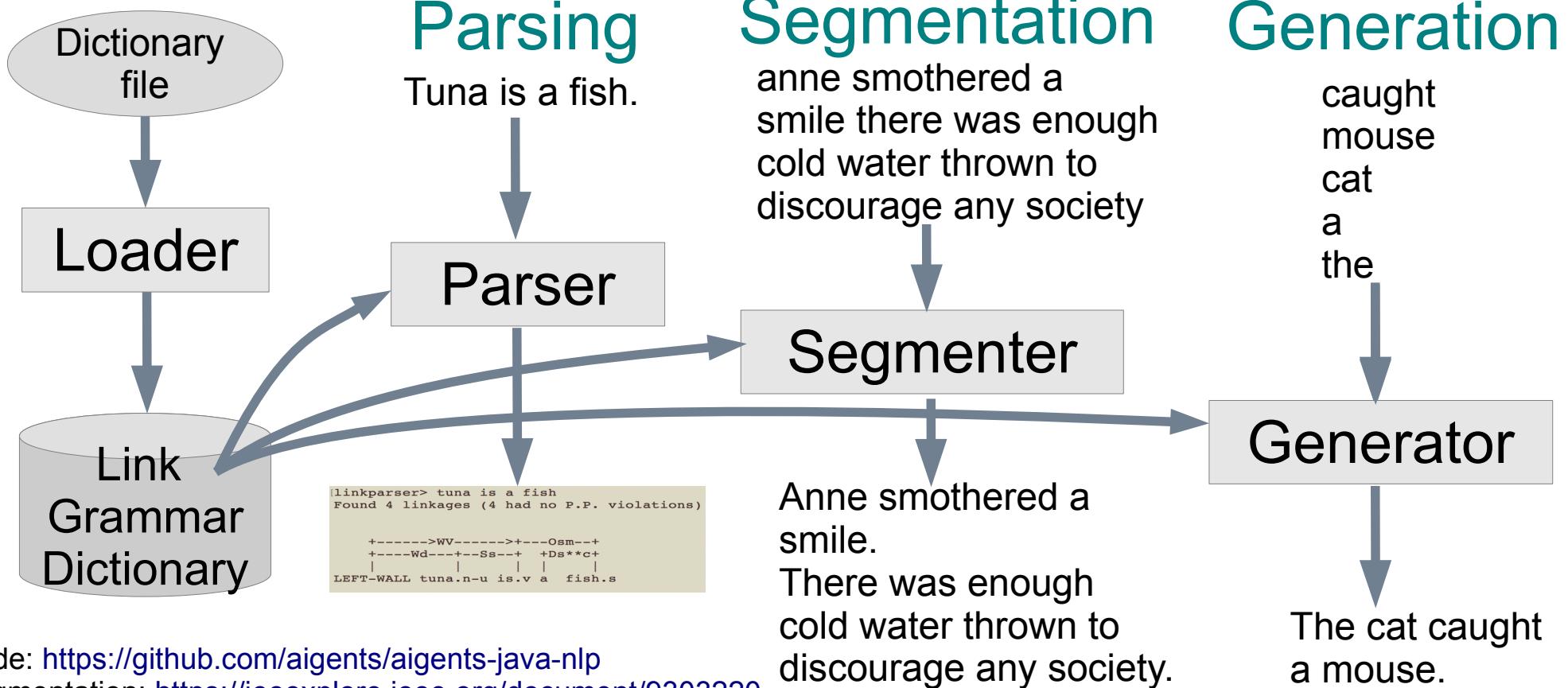
Case 3: Prediction on Cognitive Distortions for Trading



<https://arxiv.org/abs/2303.02342>

https://link.springer.com/chapter/10.1007/978-3-031-33469-6_19

Case 4: Link-Grammar based applications



Code: <https://github.com/aigents/aigents-java-nlp>

Segmentation: <https://ieeexplore.ieee.org/document/9303220>

Generation: <https://arxiv.org/abs/2105.00830>

V. Ramesh, A. Kolonin, 2020-2021

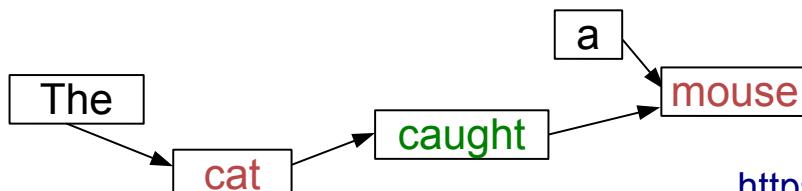
Case 4.2: Grammatical Language Generation

Generator determines what sentences can be formed from a given list of words via valid Link Grammar rules:

- 1) Given a list of words, the Generator determines a subset of all orderings of those words that satisfies initial checks of the planarity and connectivity metarules.
- 2) For each ordering in the subset, the Generator determines if that ordering is valid; specifically, it ensures that every pair of consecutive words can be connected via links part of the Dictionary objects. To do so, the Generator uses the `connects()` function, which returns a boolean value indicating whether its two parameters left and right can be linked together.

Planarity metarule: links do not cross

Connectivity metarule: links and words of a sentence must form a connected graph that can be completely traversed via one path



<https://github.com/aigents/aigents-java-nlp>

<https://arxiv.org/abs/2105.00830> V. Ramesh, A. Kolonin, 2021

Copyright © 2024 Anton Kolonin, Aigents®

Metric	Result
Single correct generated sentence	62/92
Multiple sentences with one correct	30/92
Multiple sentences with none correct	0/92
No generated sentences	0/92
Too many results*	0/92
Accuracy	1.000
Total runtime	18 min, 46 sec
Average runtime per sentence	0 min, 12 sec

Metric	Result
Single correct generated sentence	8/92
Multiple sentences with one correct	57/92
Multiple sentences with none correct	0/92
No generated sentences	0/92
Too many results*	27/92
Accuracy	0.707
Total runtime	115 min, 6 sec
Average runtime per sentence	1 min, 15 sec

* “Too many results” is defined as over 25 generated sentences.

Metric	Result
Single correct generated sentence	1/54
Multiple sentences with one correct	53/54
Multiple sentences with none correct	0/54
No generated sentences	0/54
Accuracy	1.000
Total runtime	141 min, 51 sec
Average runtime per sentence	2 min, 37 sec

Case 4.3: Question Answering with Link Grammar

Context: Identity and Relationships

A **mom** is a **human**. A **dad is a human**. A **mom** is a parent. A **dad** is a parent. A son is a child. A **daughter** is a child. A son is a **human**. A **daughter** is a **human**. **Mom** is a **human** now. **Dad** is a **human** now. **Mom** is a parent now. **Dad** is a parent now. Son is a child now. **Daughter** is a child now. Son is a **human** now. **Daughter** is a **human** now. **Mom was a daughter before.** **Dad** was a son before. **Mom** was not a parent before. **Dad** was not a parent before.

Question: mom daughter

Answer: [Mom was a daughter before., Daughter was a mom before.]

Question: dad human

Answer: [Dad is a human.]

Table 1. Results when tested on 60 queries from SingularityNET's POC-English corpus.

Metric	Results				
	Ours	BERT	ELECTRA	DistilBERT	RoBERTa
BLEU	0.878	0.639	0.712	0.604	0.767
WVCS	0.944	0.606	0.741	0.595	0.799
WER	0.645	0.924	0.550	1.095	0.150
TER	0.166	0.381	0.342	0.457	0.245

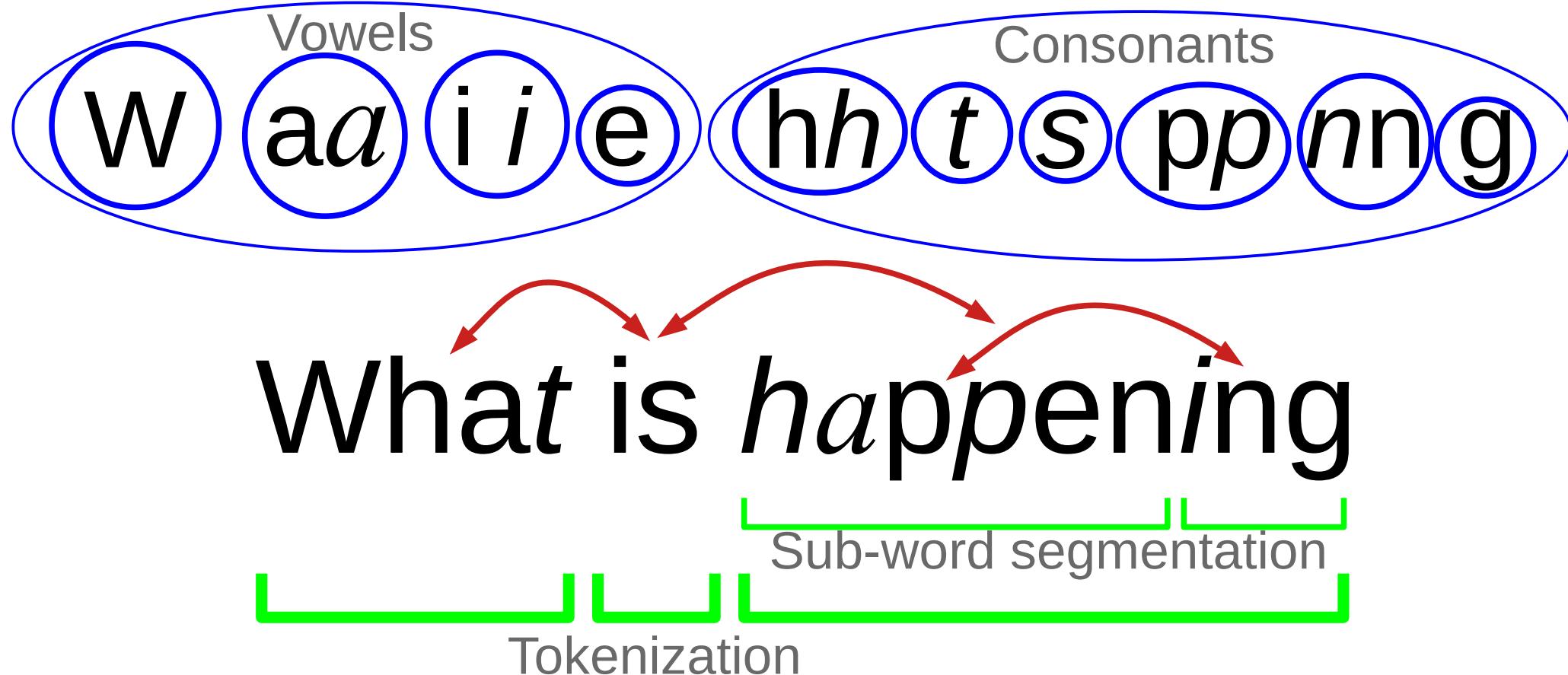
<https://www.youtube.com/watch?v=MKIOqO9FRq0>

https://link.springer.com/chapter/10.1007/978-3-030-93758-4_22 V. Ramesh, A. Kolonin, 2021

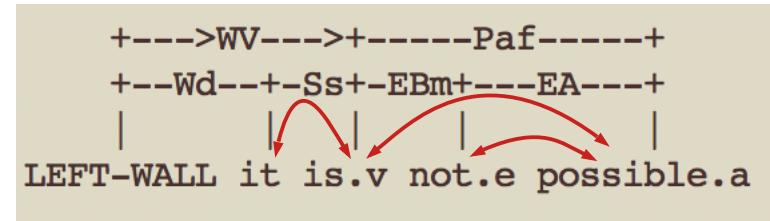
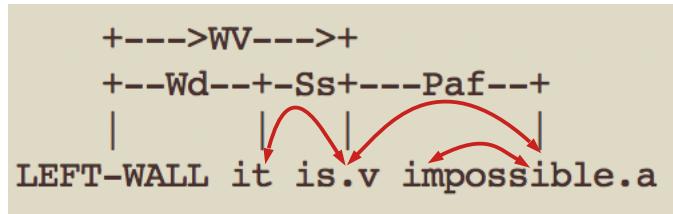
A bit more theory: Learning symbolic “models”

Learning Lexicon, Punctuation, Morphology and Grammar

Clustering, Segmentation and Parsing



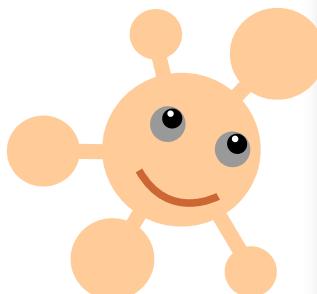
Blurring boundaries between grammatical parsing and morphological parsing causes tokenization ambiguity



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

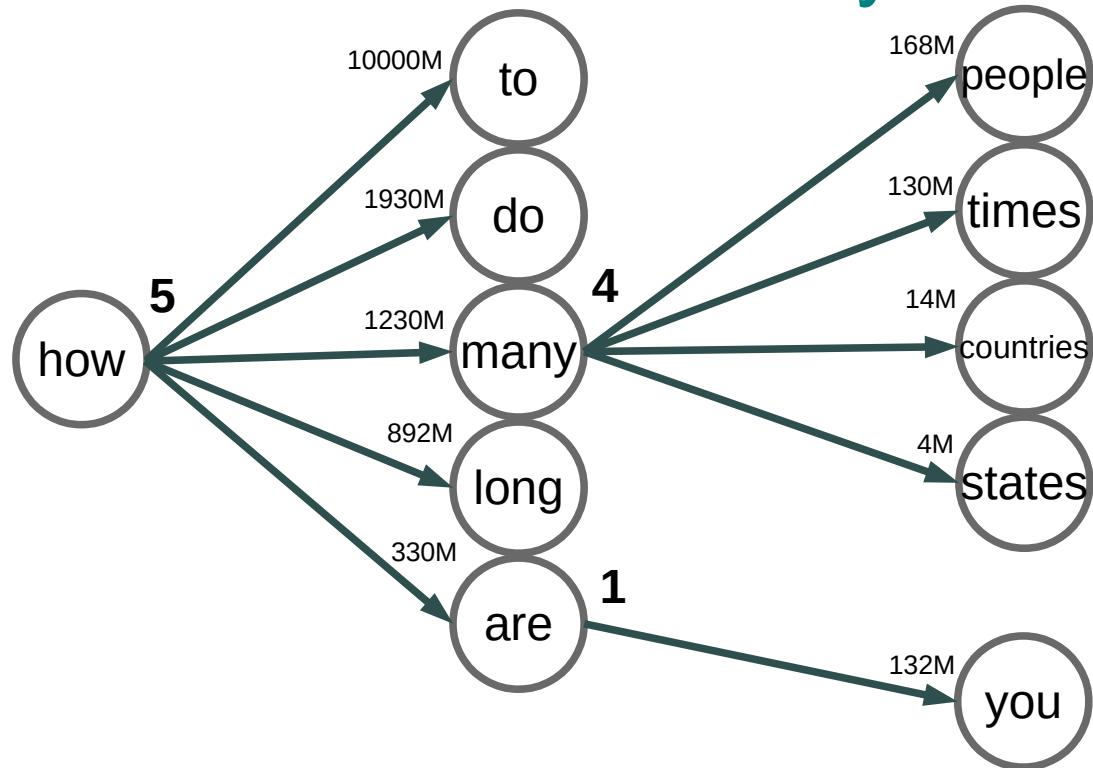
<https://aclanthology.org/2022.emnlp-main.239/>

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

Copyright © 2024 Anton Kolonin, Aigents®

21

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>

A few more cases: Learning and using symbolic “models”

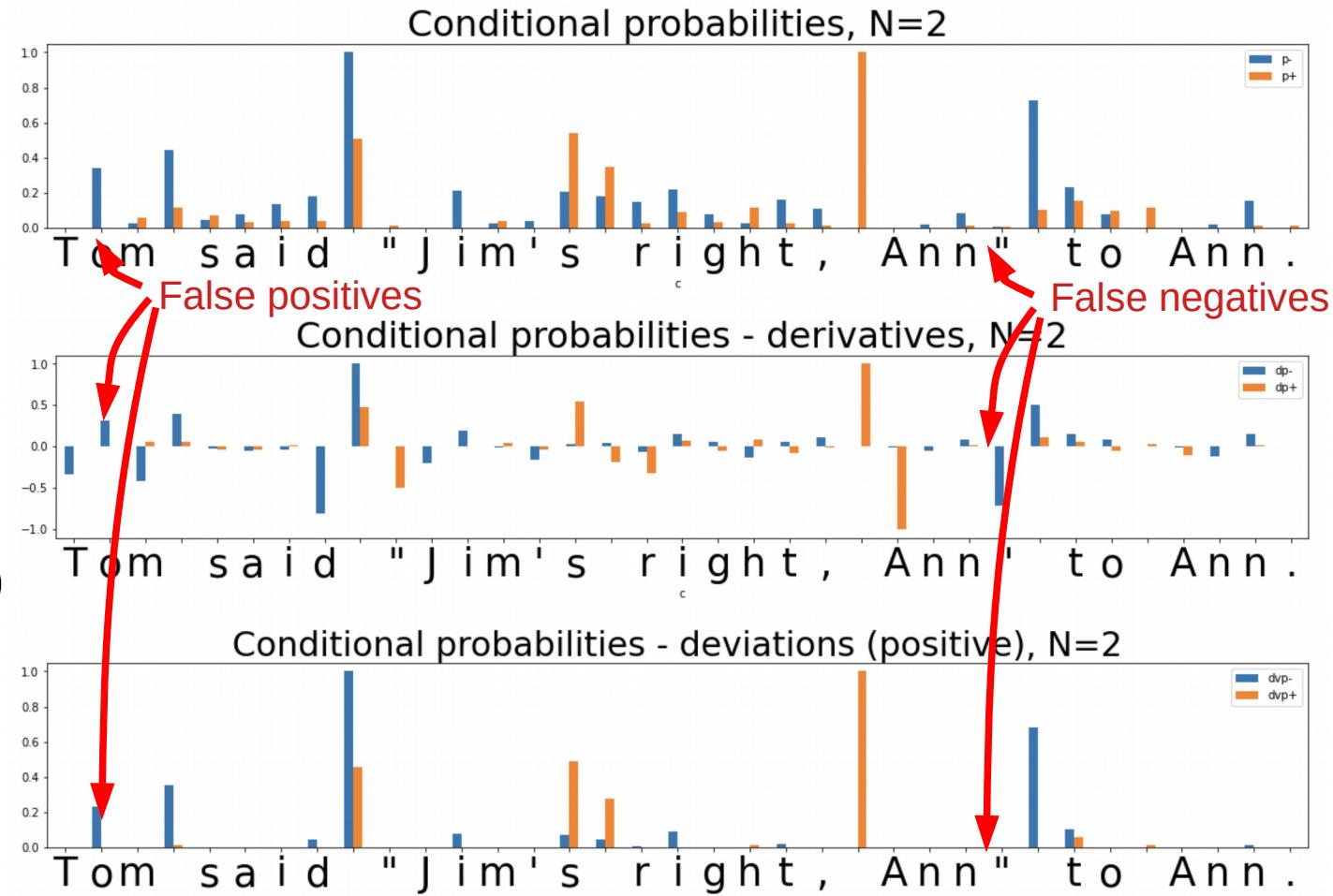
Case 5: Unsupervised Text Segmentation (Probability)

Metrics/Indicators:

Ngram (Character)
Conditional
Probability
(of Transition)

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

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



Case 5: Unsupervised Text Segmentation (Uncertainty)

Metrics/
Indicators:

Transition
Freedom
Deviation

(varying “N”
of N-gram)



Case 5: Unsupervised Text Segmentation (Hyper-parameters)

English

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

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

Hyper-Parameters:

Metric:
Transition
Freedom

Threshold
for model
compression

Combination
of Ngram N-s

Threshold for
segmentation

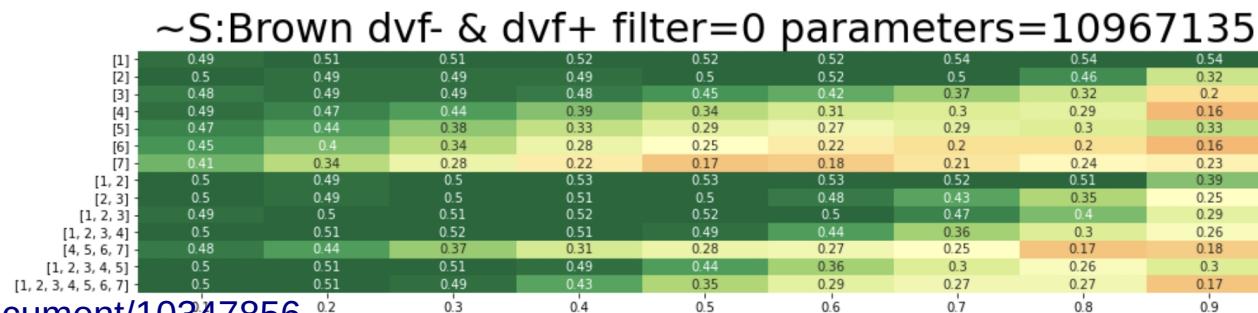
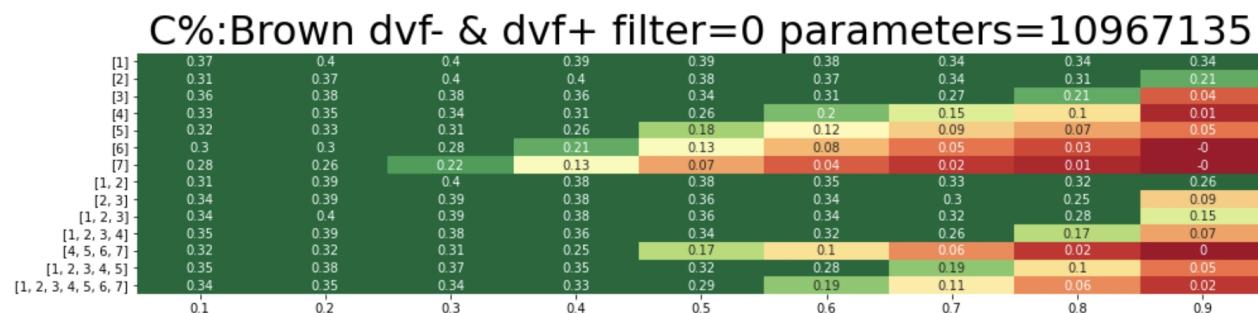
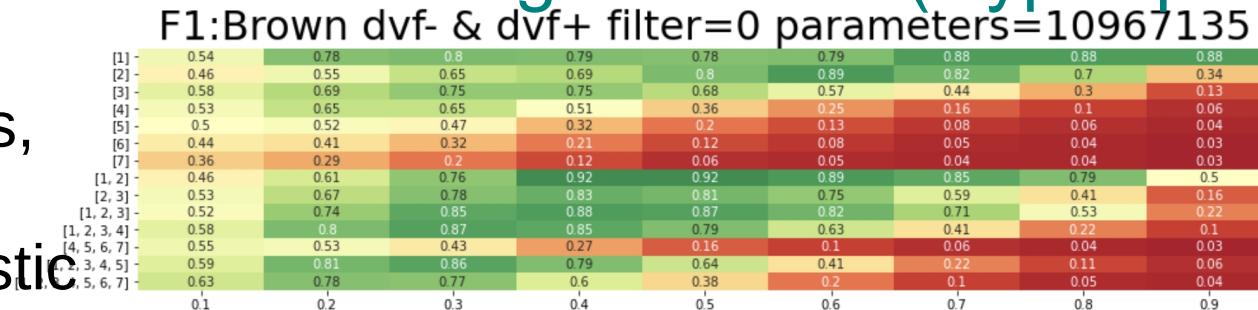
F1 - Brown ddf- & ddf+ filter=0.0001 parameters=8643703

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

Case 5: Unsupervised Text Segmentation (Hyper-parameters)

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

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

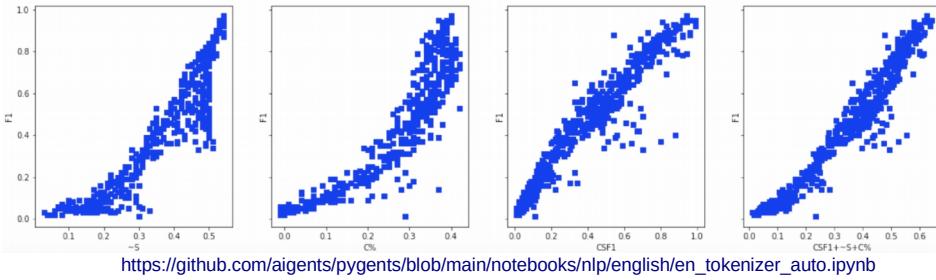


<https://ieeexplore.ieee.org/document/10347856>

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

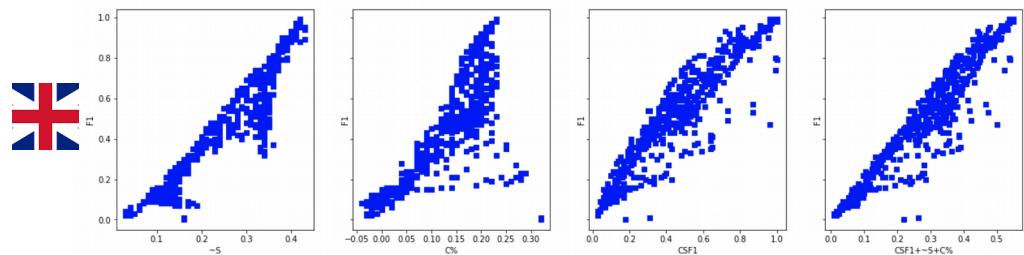
Case 5: Unsupervised Text Segmentation (Different corpora)

English, Train: Brown, Test: Brown 1000

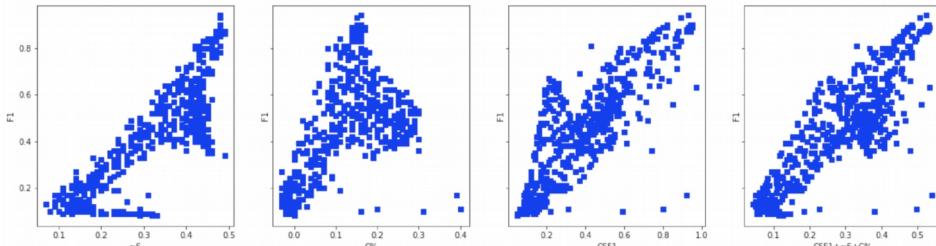


https://github.com/agents/pygents/blob/main/notebooks/nlp/english/en_tokenizer_auto.ipynb

English, Train: Brown, Test: MagicData 100

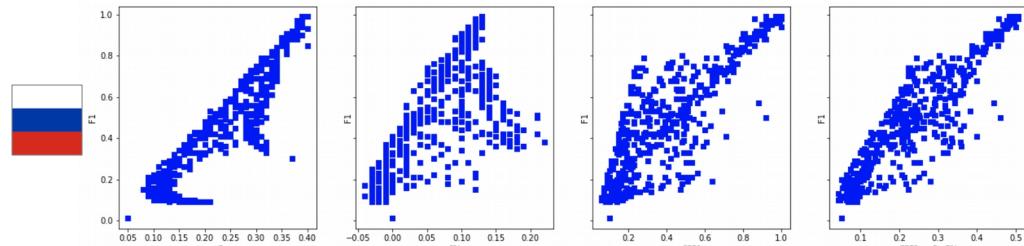


Russian, Train: RusAge, Test: RusAge 1000

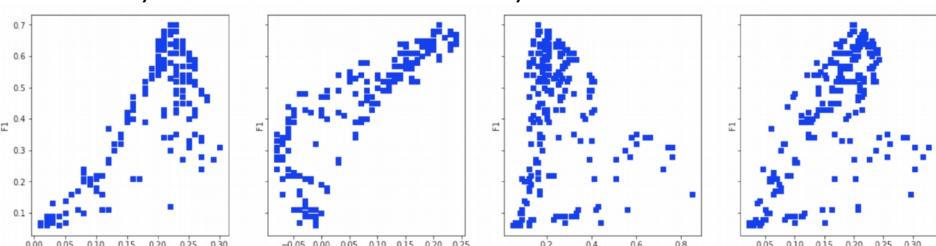


https://github.com/agents/pygents/blob/main/notebooks/nlp/russian/ru_tokenizer_auto.ipynb

Russian, Train: Brown, Test: MagicData 100

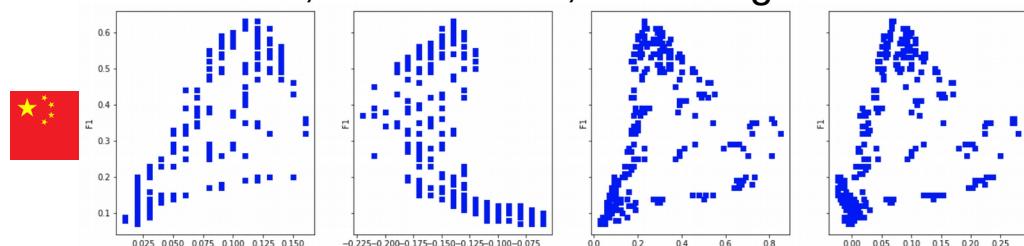


Chinese, Train: CLUE News, Test: CLUE News 1000



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 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

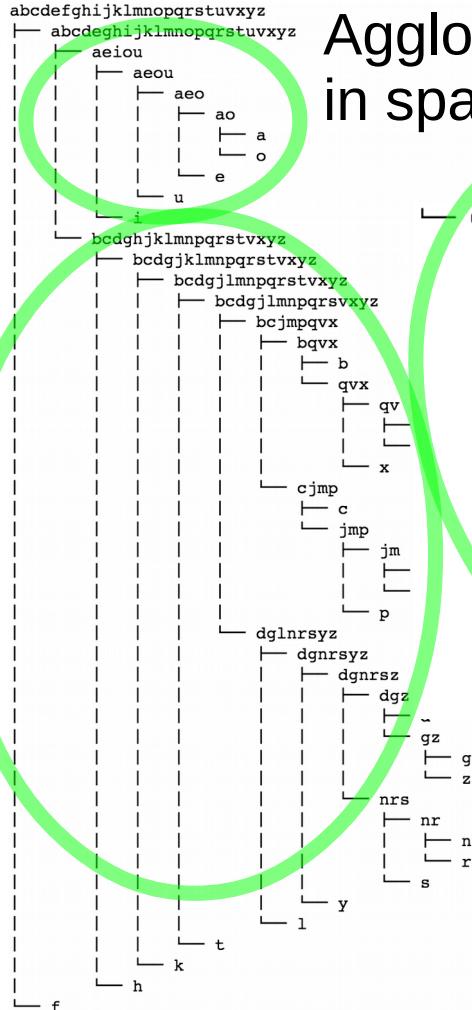
<https://aclanthology.org/2022.emnlp-main.239/>

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

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

Case 6: Unsupervised Character Category Learning

Agglomerative Clustering in space of Transitions

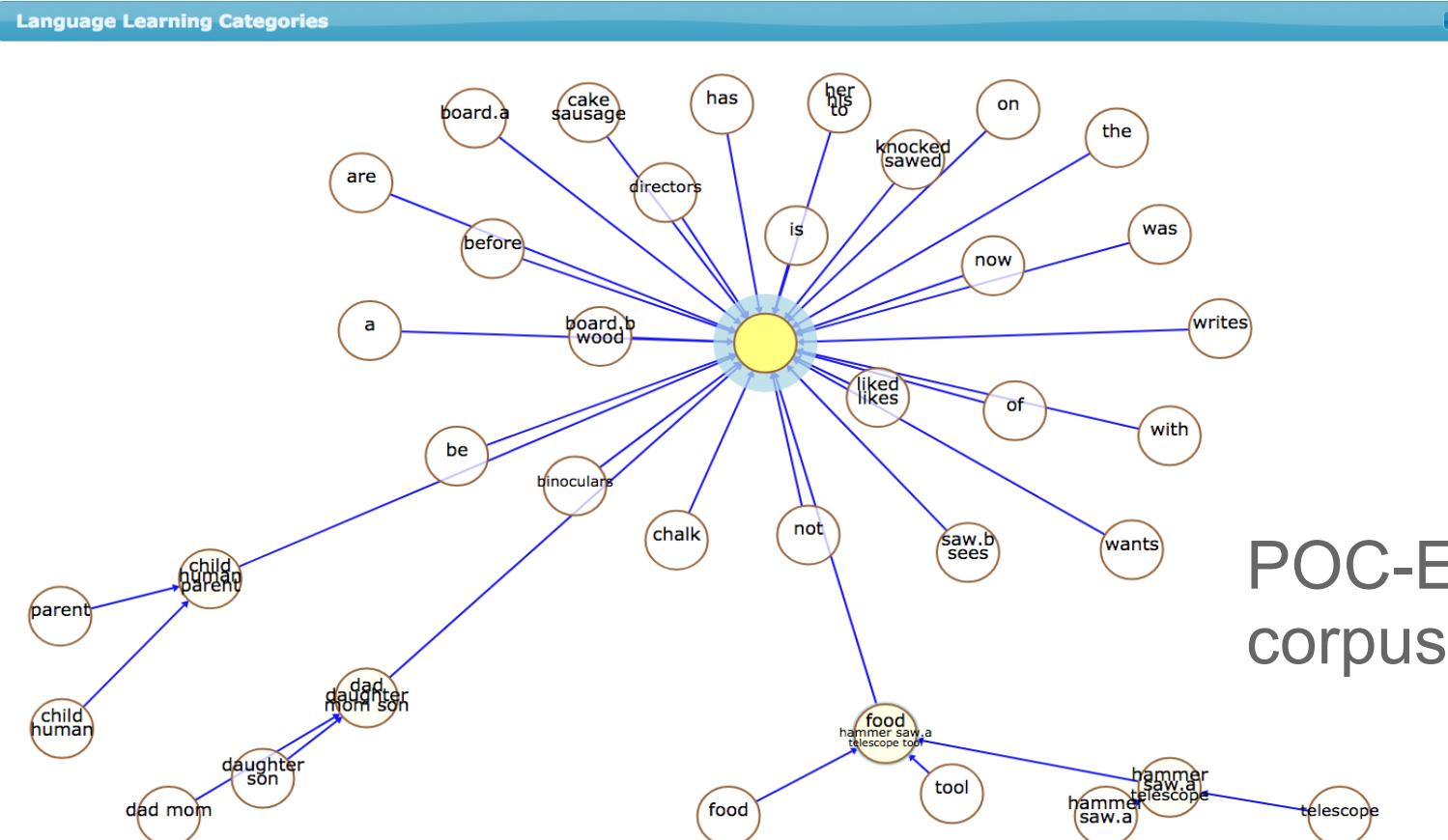


RusAge Test/Small,
Cosine Similarity

Copyright © 2024 Anton Kolonin, Agents®

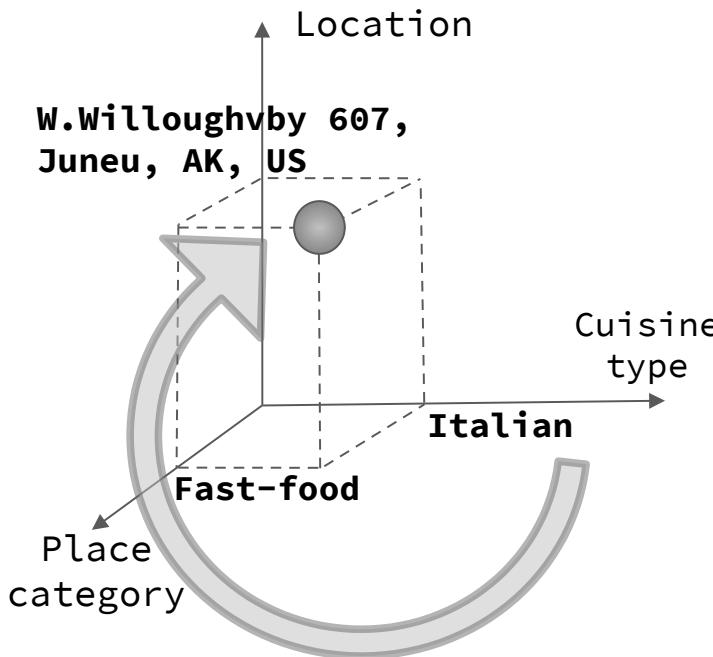
<https://aclanthology.org/2022.emnlp-main.239.pdf>

Case 7: Unsupervised Ontology Learning on Parses



POC-English corpus

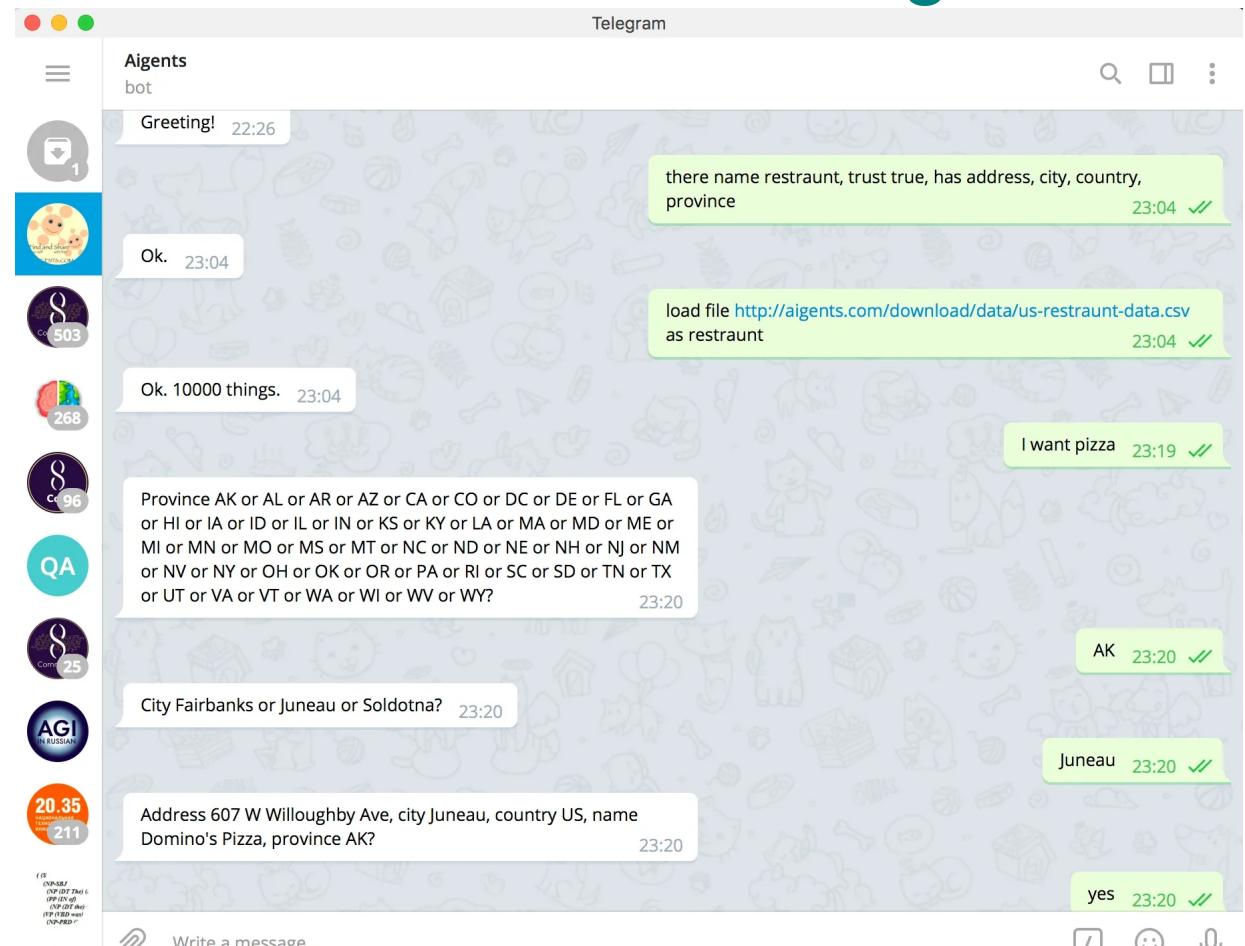
Case 8: Data-driven Conversational Intelligence



Dynamic data-driven dialog
flow minimizing uncertainty
Just add RAG augmentation?

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

Copyright © 2024 Anton Kolonin, Aigents®



Welcome to the Interpretable Natural Language Processing Community and Series of Workshops

Telegram English
<https://t.me/internlp>

Telegram Russian
<https://t.me/agibots>

INLP Workshops
<https://agents.github.io/inlp/>

Extended presentation in Russian
<https://www.youtube.com/watch?v=RU2T0OFJJB4>

Thank You and Welcome!



<https://agirussia.org>

Anton Kolonin
akolonin@aigents.com
Messengers: akolonin

