Transparent information extraction from natural language

Gábor Recski

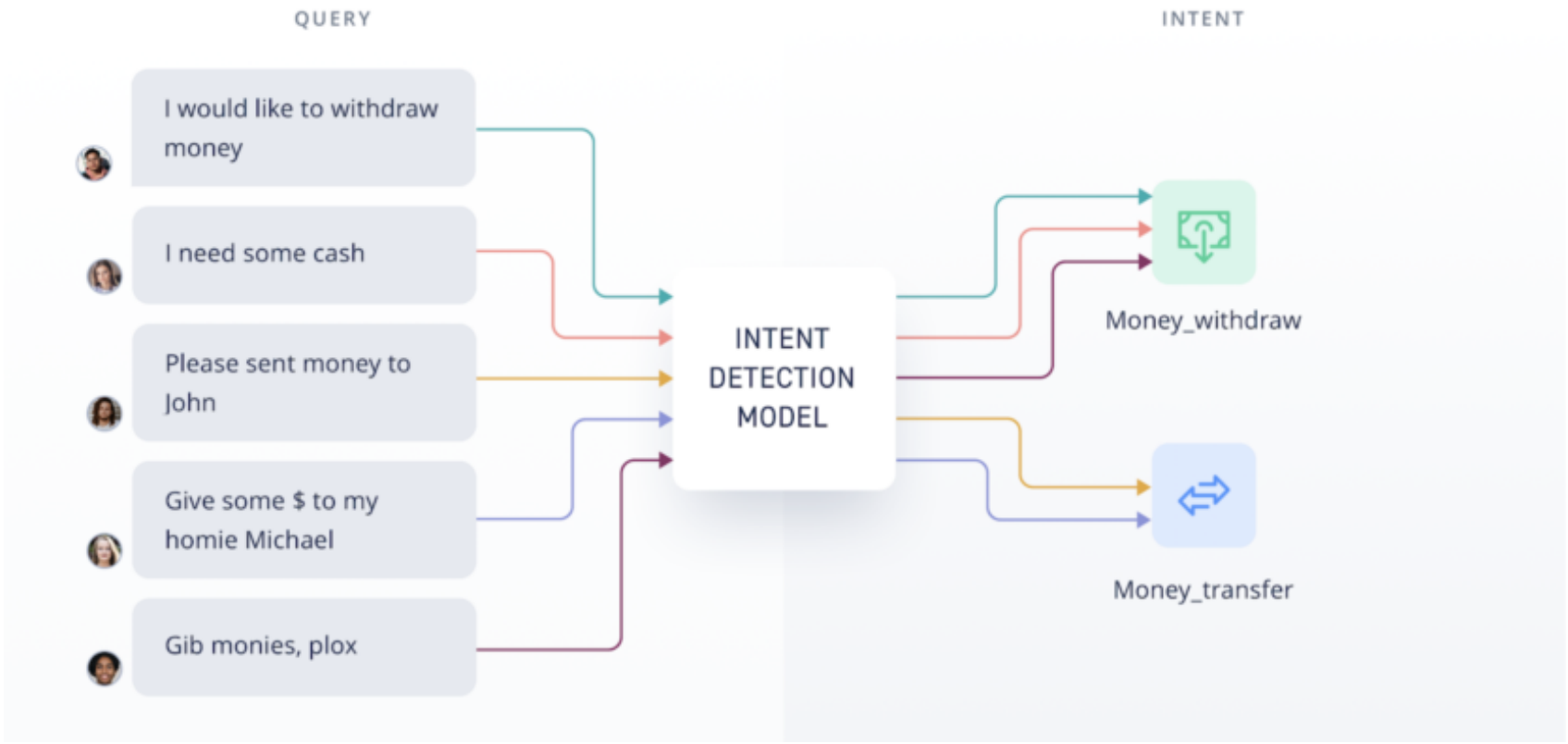
TU Wien

gabor.recski@tuwien.ac.at

CSH seminar
06/22/2022

Information extraction (IE)

IE example: Text classification



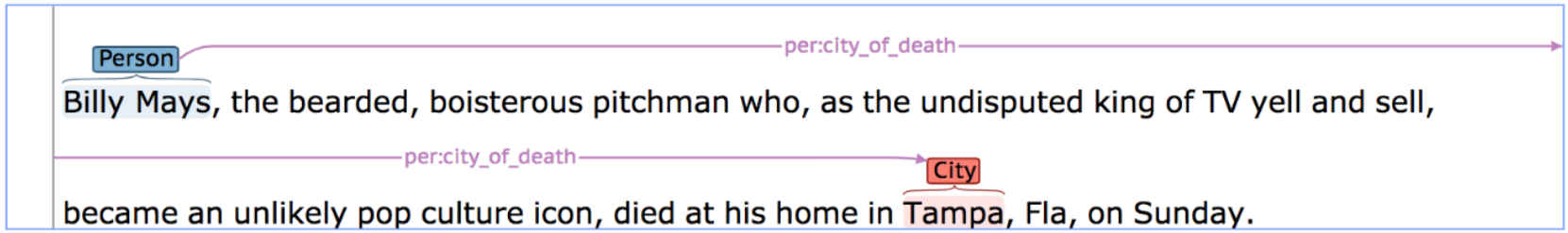
[9]

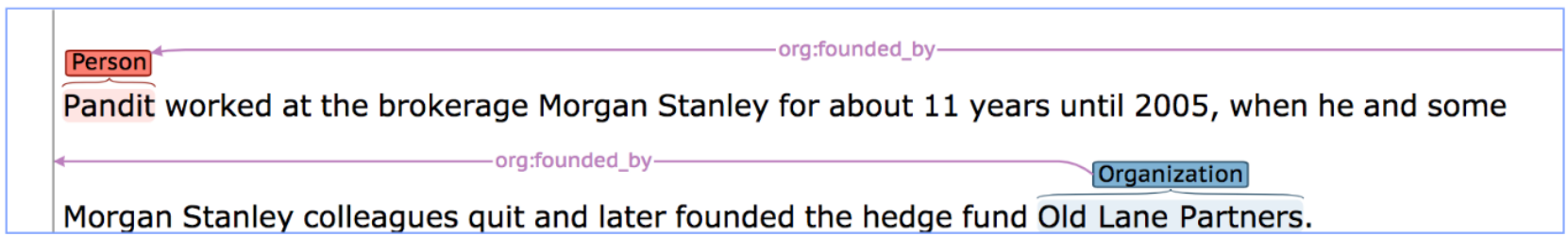
IE example: Named Entity Recognition (NER)

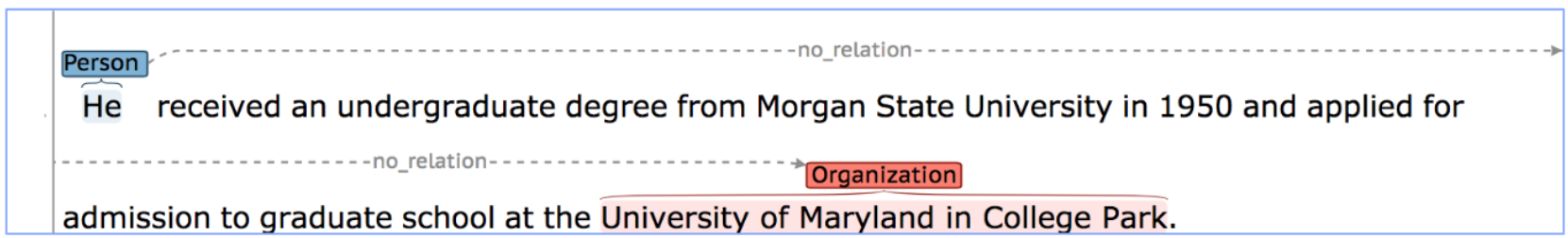
Citing high fuel prices, [$_{\text{ORG}}$ **United Airlines**] said [$_{\text{TIME}}$ **Friday**] it has increased fares by [$_{\text{MONEY}}$ **$6**] per round trip on flights to some cities also served by lower-cost carriers. [$_{\text{ORG}}$ **American Airlines**], a unit of [$_{\text{ORG}}$ **AMR Corp.**], immediately matched the move, spokesman [$_{\text{PER}}$ **Tim Wagner**] said. [$_{\text{ORG}}$ **United**], a unit of [$_{\text{ORG}}$ **UAL Corp.**], said the increase took effect [$_{\text{TIME}}$ **Thursday**] and applies to most routes where it competes against discount carriers, such as [$_{\text{LOC}}$ **Chicago**] to [$_{\text{LOC}}$ **Dallas**] and [$_{\text{LOC}}$ **Denver**] to [$_{\text{LOC}}$ **San Francisco**].

[4, Ch. 22]

IE example: Relation extraction

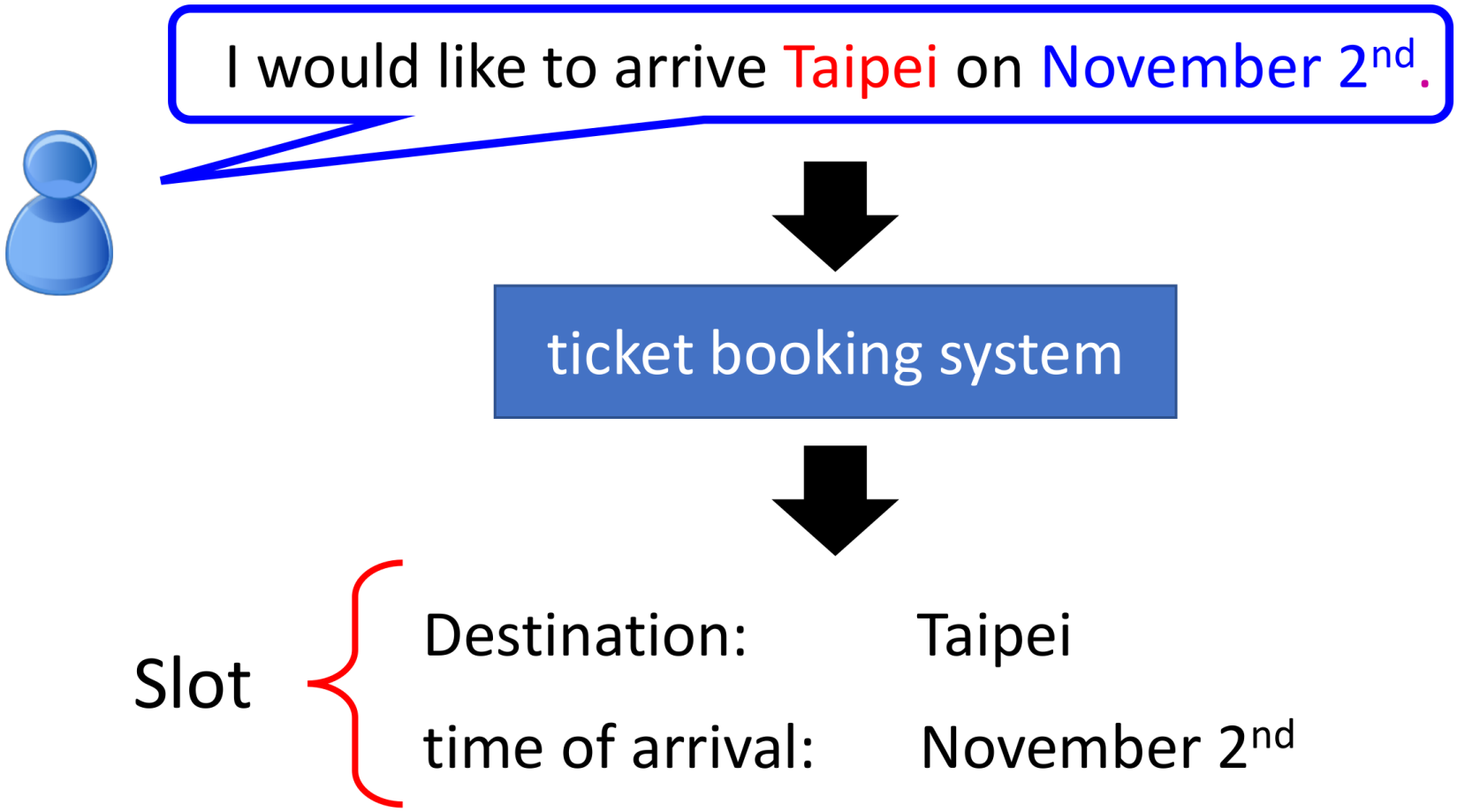






[15]

IE example: Slot filling



[8]

IE approaches: machine learning

IE approaches: machine learning

Process:

- annotate training data
- train model
- if solution is inadequate, blame the data and/or the user
- repeat (until funding runs out)

IE approaches: machine learning

Process:

- annotate training data
- train model
- if solution is inadequate, blame the data and/or the user
- repeat (until funding runs out)

Why are ML models inadequate?

IE approaches: machine learning

Process:

- annotate training data
- train model
- if solution is inadequate, blame the data and/or the user
- repeat (until funding runs out)

Why are ML models inadequate?

- learns patterns, not tasks
- learns artefacts/unintended bias [1, 3, 10, 13, 14, 7, 11]
- not configurable
- limited/false explainability [2, 11]
- black box (even for developers!)

IE approaches: rule-based systems

- explainable, interpretable
- flexible, customizable
- reliable
- auditable

IE approaches: rule-based systems

- explainable, interpretable
- flexible, customizable
- reliable
- auditable

Why isn't everyone using them?

IE approaches: rule-based systems

- explainable, interpretable
- flexible, customizable
- reliable
- auditable

Why isn't everyone using them?

- they are, actually...
- expensive to build and maintain/reconfigure
- requires both domain and tech expertise
- **need to model variability of natural language for each task**

Our approach: semantics-based IE

Step 1: map text to a generic intermediate (semantic) representation

- task-independent
- domain-independent
- syntax-independent
- language-independent*

Step 2: map semantic representation to task-specific structures

- as specific as necessary
- intuitive model of the task's semantics

People involved


Kinga Gémes

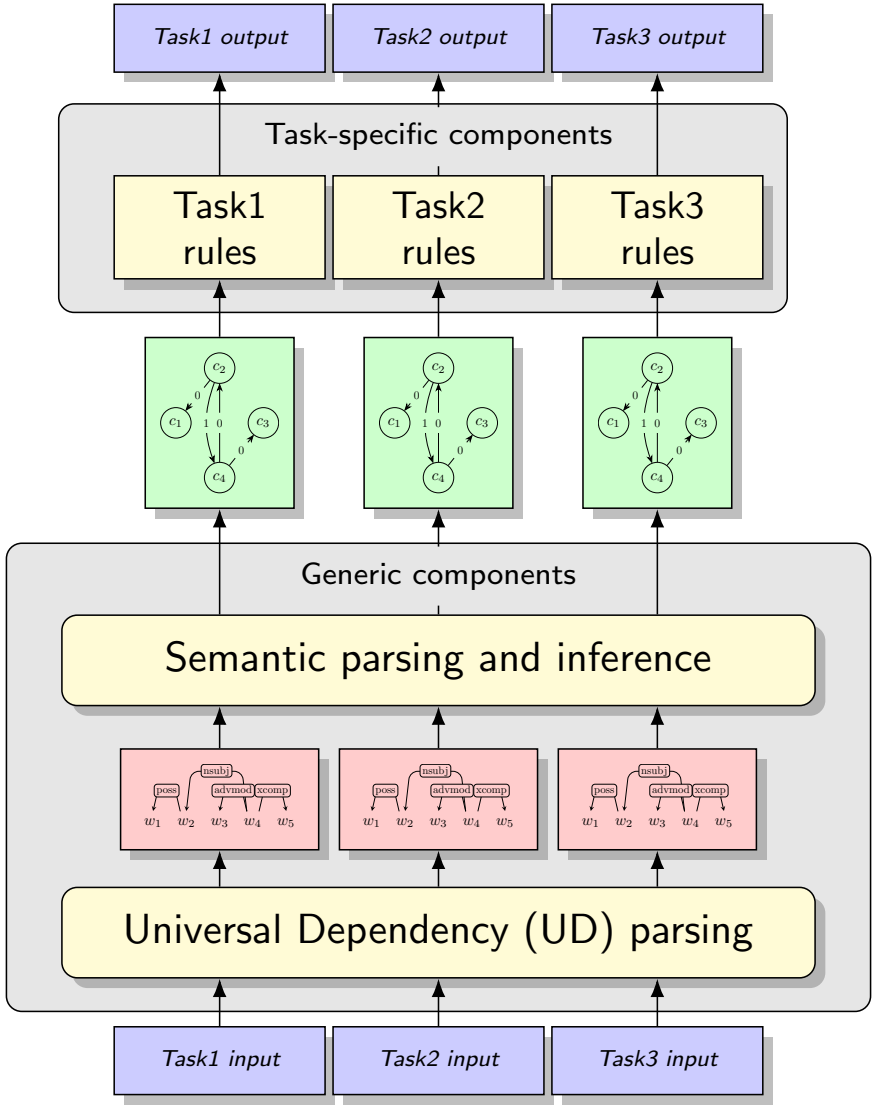

Eszter Iklódi


Ádám Kovács


Gábor Recski

https://nlp.ec.tuwien.ac.at/people/

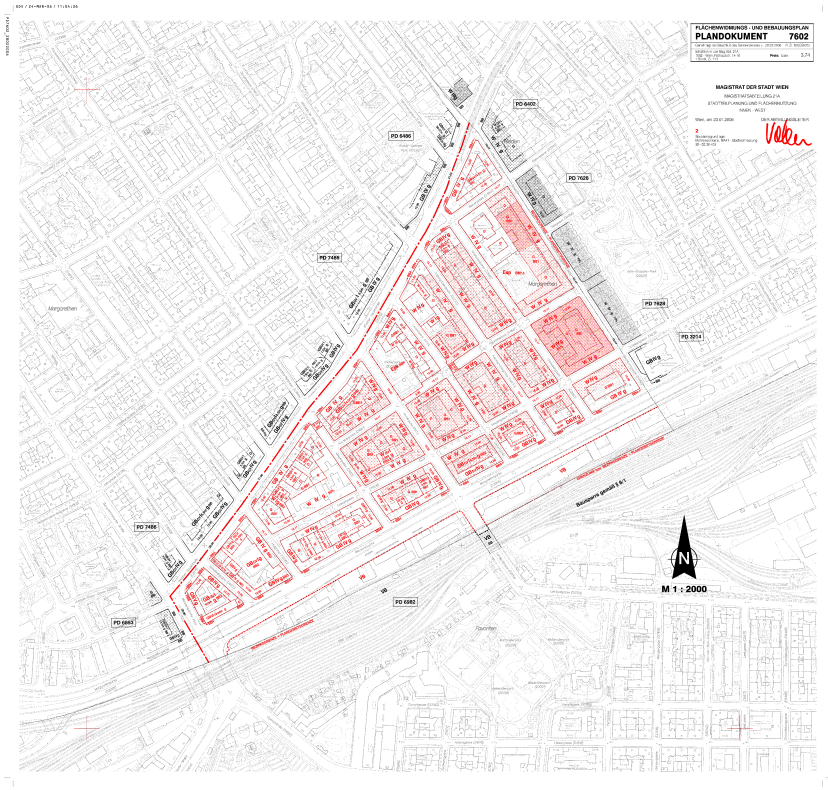
Our approach: semantics-based IE

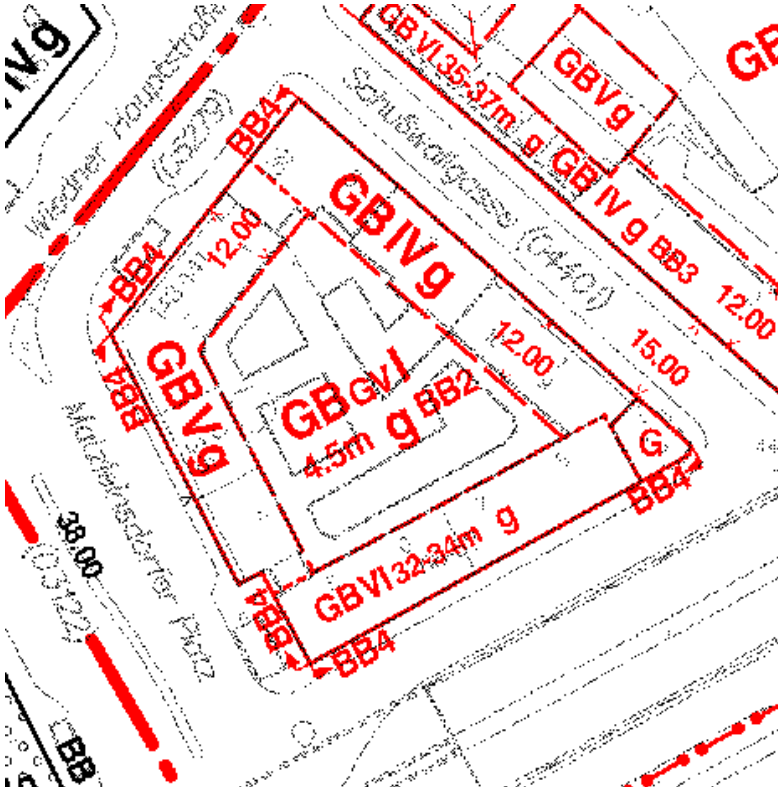


The BRISE use-case

The BRISE use-case

Zoning plans (Flächenwidmungs- und Bebauungspläne)





The BRISE use-case

Textual regulations (Textliche Bestimmungen)

4. 1. Auf den mit **BB1** bezeichneten Grundflächen ist die Errichtung von unterirdischen und oberirdischen Gebäuden untersagt.
4. 2. Auf den mit **BB2** bezeichneten Grundflächen sind die zur Errichtung gelangenden Dächer entsprechend dem Stand der Technik als begrünte Flachdächer auszuführen.
4. 3. Auf den mit **BB3** bezeichneten Grundflächen ist die Errichtung von Staffelgeschossen an den Straßenfronten zulässig.
4. 4. Entlang der mit **BB4** bezeichneten Baulinien dürfen keine Hauptfenster von Aufenthaltsräumen von Wohnungen im Erdgeschoss zu den Verkehrsflächen hin orientiert werden.
4. 5. Auf den mit **BB5** bezeichneten und als Grünland/Erholungsgebiet/Sport- und Spielplätze gewidmeten Grundflächen dürfen Gebäude bis zu einer Gebäudehöhe von 4,5 m errichtet werden. Das Ausmaß der bebauten Fläche darf maximal 10 v.H. der gesamten Grundfläche betragen.
4. 6. Auf den mit **BB7** bezeichneten Grundflächen ist in der Höhenlage der anschließenden Verkehrsfläche eine öffentliche Durchfahrt mit einer lichten Höhe von 3,0 m freizuhalten und zu dulden.

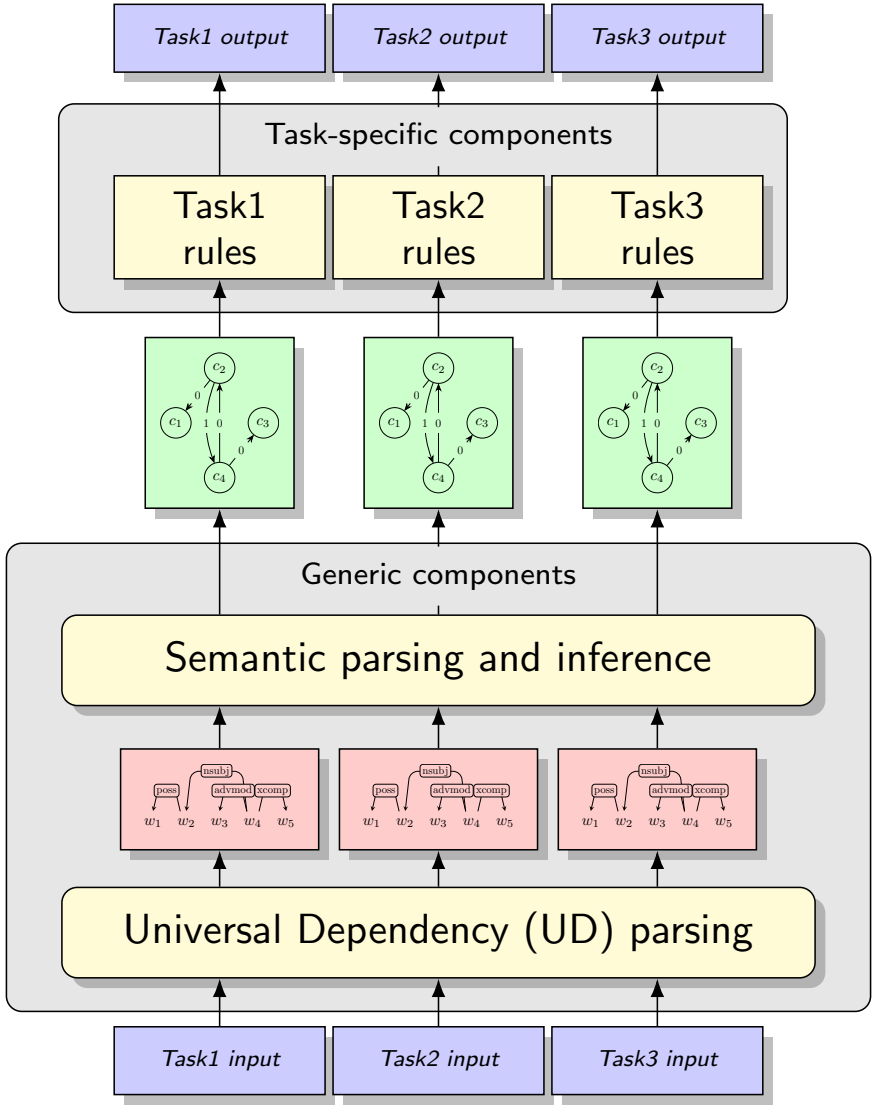
The BRISE rule extraction task

Flachdächer bis zu einer Dachneigung von fünf Grad sind entsprechend dem Stand der technischen Wissenschaften zu begrünen.

Flat roofs with a pitch not exceeding 5 degrees must be greened using state of the art technologies.

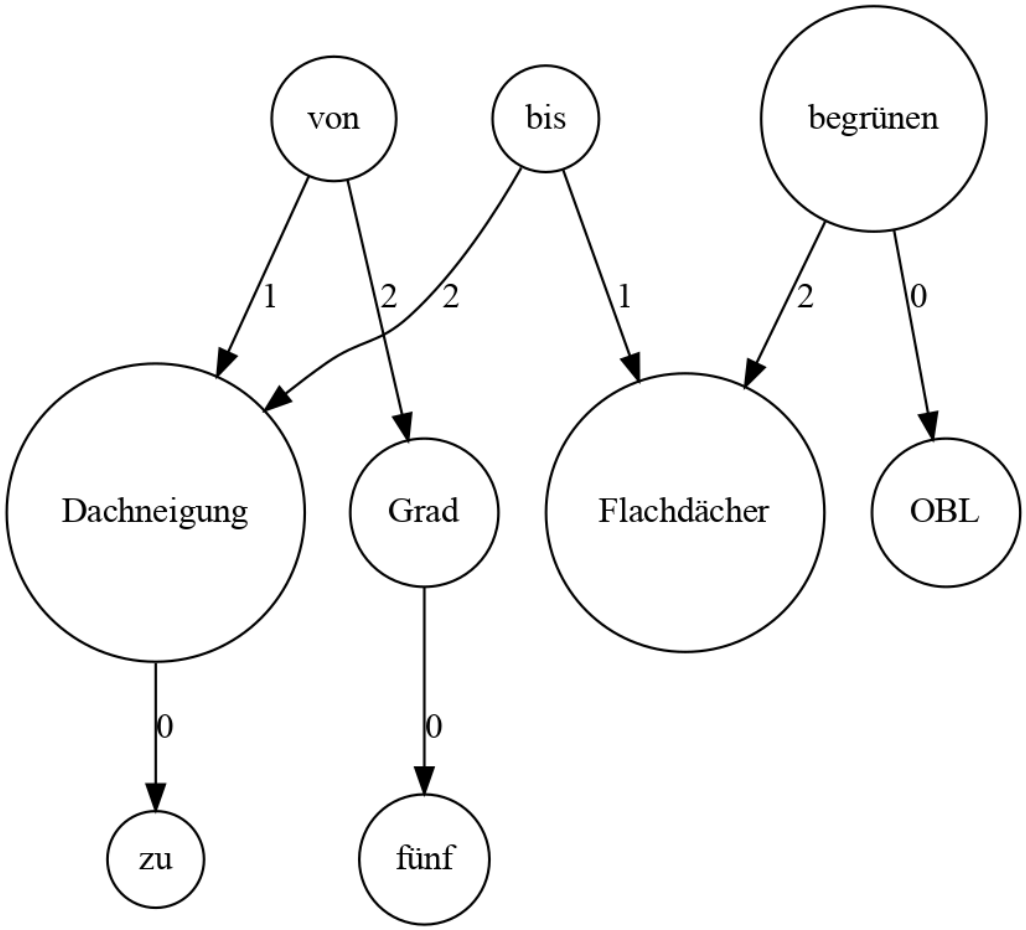
```
{"modality": "obligation",
 "attributes": [
    {"type": "content",
     "name": "BegruenungDach",
     "value": true},

    {"type": "condition",
     "name": "Dachart",
     "value": "Flachdach"},

    {"type": "condition",
     "name": "DachneigungMax",
     "value": "5Grad"}]}
```

Semantics-based IE



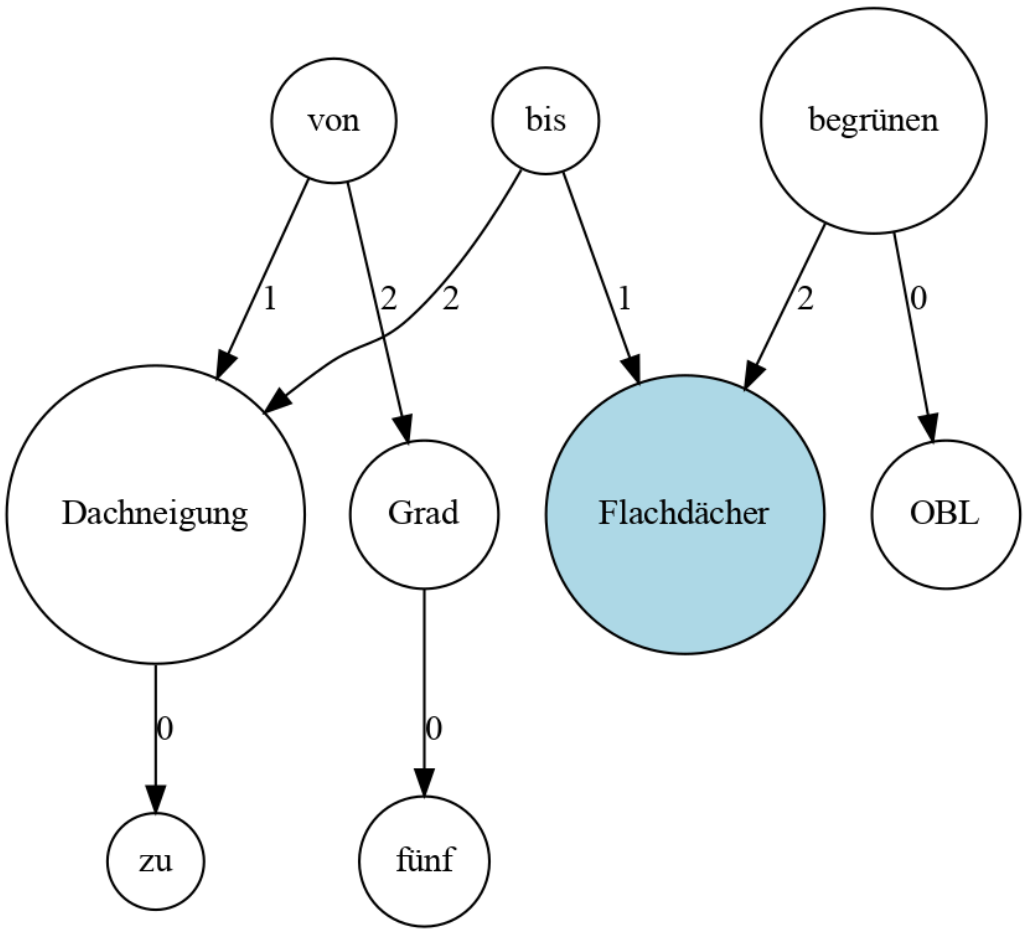
Extraction via graph patterns

Flachdächer bis zu einer Dachneigung von fünf Grad sind . . . zu begrünen.



Extraction via graph patterns

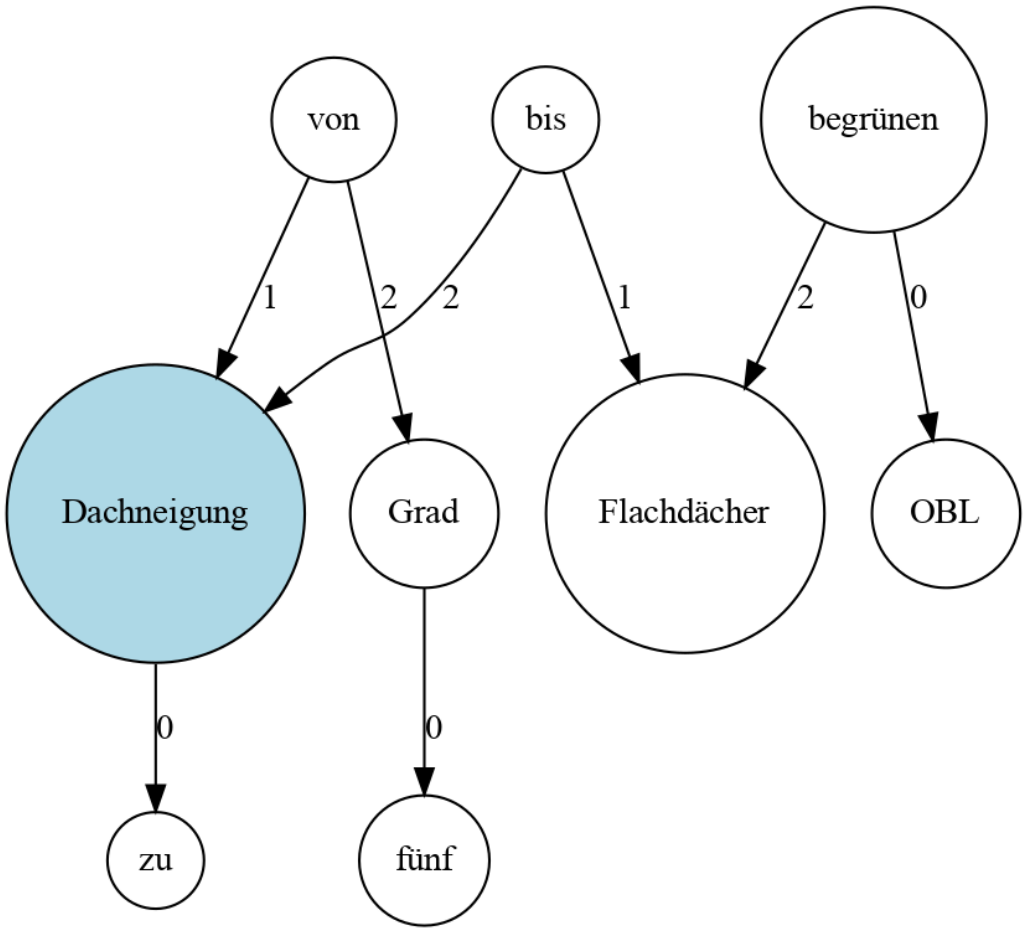
Flachdächer bis zu einer Dachneigung von fünf Grad sind . . . zu begrünen.



`Dachart`

Extraction via graph patterns

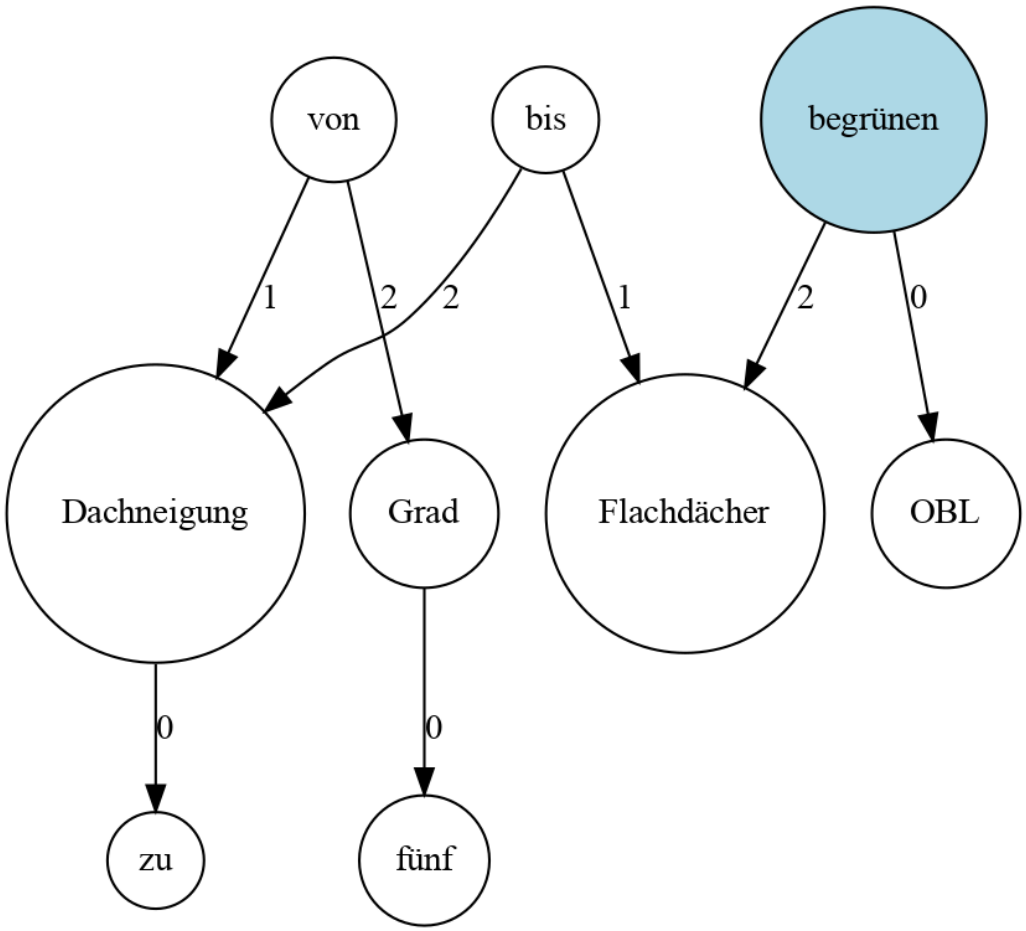
Flachdächer bis zu einer Dachneigung von fünf Grad sind . . . zu begrünen.



`DachneigungMax`

Extraction via graph patterns

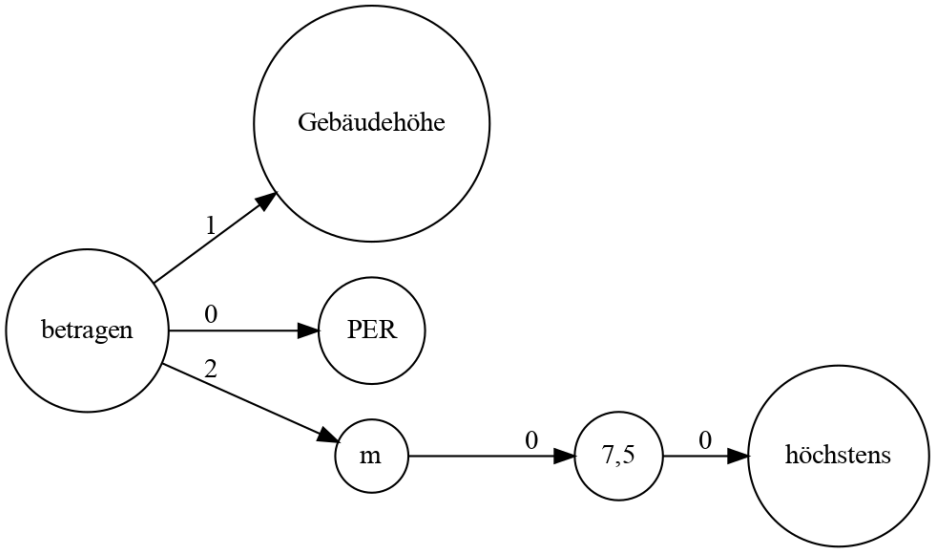
Flachdächer bis zu einer Dachneigung von fünf Grad sind . . . zu begrünen.



`BegrueungDach`

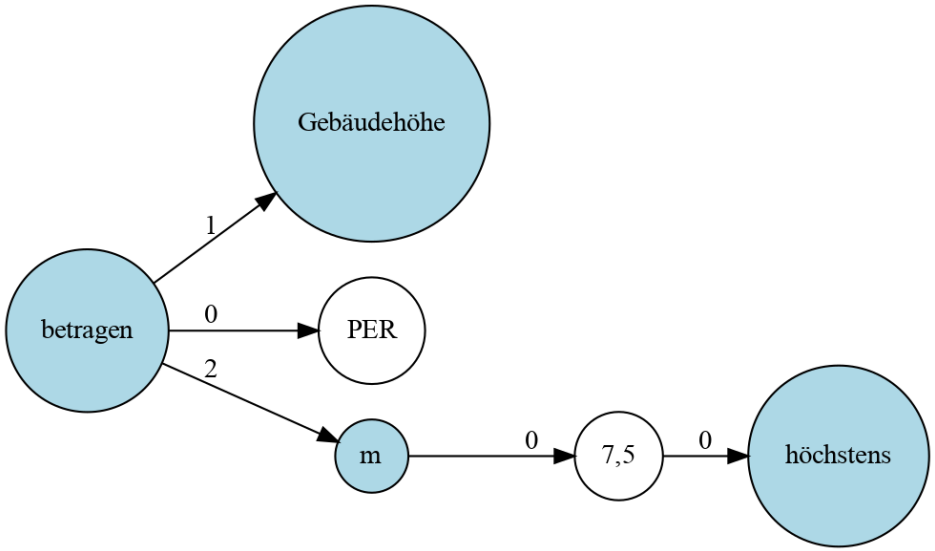
Extraction via graph patterns

Die Gebäudehöhe darf höchstens 7,5 m betragen.



Extraction via graph patterns

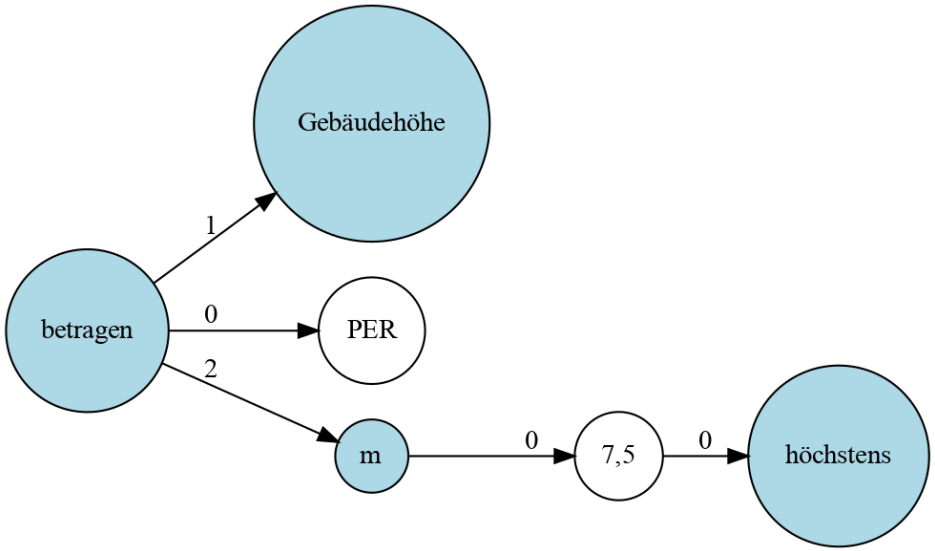
Die Gebäudehöhe darf höchstens 7,5 m betragen.



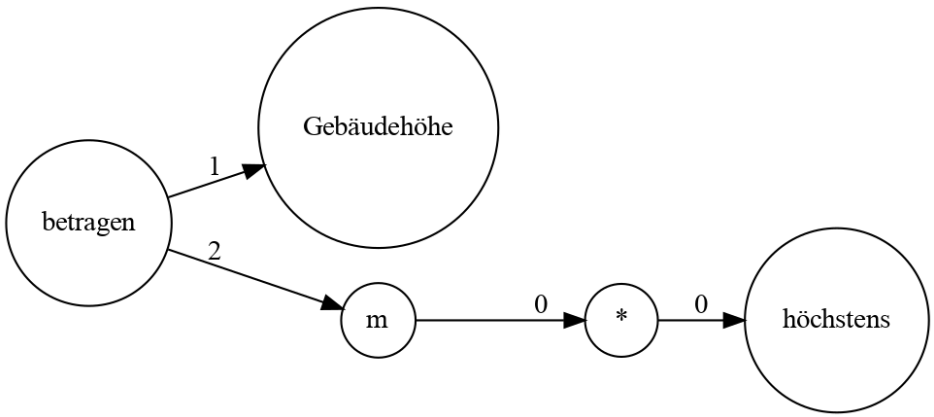
`GebaeudeHoeheMax`

Extraction via graph patterns

Die Gebäudehöhe darf höchstens 7,5 m betragen.



`GebaeudeHoeheMax`



But how can we build such rules efficiently?
And without much technical background?

- pure Python library for building rule-based text classifiers over graphs
- interactive UI for human-in-the-loop (HITL) learning of graph patterns
- supports multiple linguistic graph formalisms (UD, AMR, stanza)
- provides a REST-API to use extracted features for inference in production mode
- open-source (MIT license), installable via pip

HITL learning

Rule chooser and modifier

First, choose class you want to use to build rules

GebaeudeHoeheMax

You can modify any rule you want to

Remember, we use the PENMAN notation to describe a rule. You can find more information about the rules in the README of our repository.

rules	negated_rules
(u_123 / Gebaeudehoehe)	

Graph viewer and evaluator

Browse graphs: +

Choose from the rules

(u_123 / Gebaeudehoehe)

Result of using all the rules: Precision: **0.488**, Recall: **0.824**, Fscore: **0.613**

The rule's result: Precision: **0.488**, Recall: **0.824**, Fscore: **0.613**, True positives: **42**, False positives: **44**

HITL learning

Rule chooser and modifier

First, choose class you want to use to build rules

GebaeudeHoeheMax

You can modify any rule you want to

Remember, we use the PENMAN notation to describe a rule. You can find more information about the rules in the README of our repository.

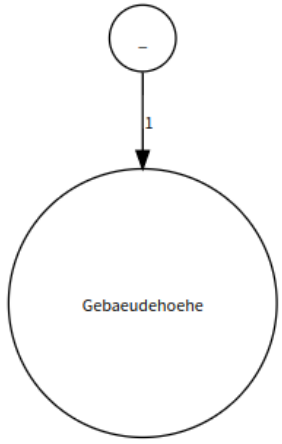
rules	negated_rules
(u_123 / Gebaeudehoehe)	
(u_1 / .* :1 (u_123 / Gebaeudehoehe))	

Graph viewer and evaluator

Browse graphs: +

Choose from the rules

(u_1 / .* :1 (u_123 / Gebaeudehoehe))



Result of using all the rules: Precision: **0.808**, Recall: **0.824**, Fscore: **0.816**

The rule's result: Precision: **0.808**, Recall: **0.824**, Fscore: **0.816**, True positives: **42**, False positives: **10**

HITL learning

You can modify any rule you want to

Remember, we use the PENMAN notation to describe a rule. You can find more information about the rules in the README of our repository.

rules	negated_rules
☐ (u_1 / .* :1 (u_123 / Gebaeudehoehe))	

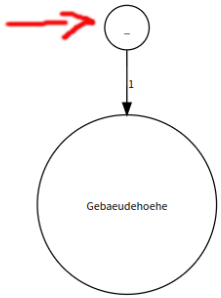
After you modified any rule, click on **save updates** button to save your changes.

Tick the **box** next to the rules you want to evaluate and then, click on the **evaluate** button. If you don't tick any rule, the evaluator will evaluate all the rules.

Delete or Refine selected rules (you can only refine underspecified rules that contain a regexp)

- none
- delete
- refine

(u_1 / .* :1 (u_123 / Gebaeudehoehe))



Result of using all the rules: Precision: **0.808**, Recall: **0.824**, Fscore: **0.816**

The rule's result: Precision: **0.808**, Recall: **0.824**, Fscore: **0.816**, True positives: **42**, False positives: **10**

Show validation data +

Select the graphs you want to view

True Positive graphs

Tick the box next to the graphs you want to see. The rule that applied will be highlighted in the graph.

The penman format of the graph will be also shown, you can copy any of the part directly from the penman format if you want to add a new rule.

id	sentence

HITL learning

rules	r
☐ (u_123 / Gebaeudehoehe)	
☐ (u_1 / .* :1 (u_123 / Gebaeudehoehe))	
☐ (u_1 / betragen|COORD|entsprechend|bis|von|ueberschreiten :1 (u_123 / Gebaeudehoehe))	

The rule's result: Precision: **0.977**, Recall: **0.824**, Fscore: **0.894**, True positives: **42**, False positives: **1**

HITL learning

rules
☐ (u_1 / betragen|COORD|entsprechend|bis|von|ueberschreiten :1 (u_123 / Gebaeudehoehe))
☐ (u_1 / betragen|bis|ueberschreiten :1 (u_123 / Gebaeudehoehe))

The rule's result: Precision: **1.000**, Recall: **0.412**, Fscore: **0.583**, True positives: **21**, False positives: **0**

HITL learning

...

HITL learning

Rule chooser and modifier

First, choose class you want to use to build rules

GebaeudeHoeheMax

You can modify any rule you want to

Remember, we use the PENMAN notation to describe a rule. You can find more information about the rules in the README of our repository.

rules	negated_rul
(u_0 / Gebaeudehoehe :0 (u_1 / maximal))	
(u_0 / ueberschreiten :1 (u_1 / Gebaeudehoehe))	
(u_0 / bis :2 (u_1 / Gebaeudehoehe))	
(u_482 / betragen :2 (u_96 / m :0 (u_744 / .* :0 (u_127 / hoechstens))) :1 (u_123 / Gebaeudehoehe))	
(u_150 / betragen :1 (u_123 / Gebaeudehoehe) :0 (u_127 / hoechstens))	
(u_482 / betragen :2 (u_96 / m :0 (u_744 / .* :0 (u_127 / maximal))) :1 (u_123 / Gebaeudehoehe))	
(u_889 / ueberschreiten :2 (u_269 / Hoehe) :1 (u_897 / Bauteil :0 (u_411 / Gebaeude)))	

After you modified any rule, click on **save updates** button to save your changes.

Tick the **box** next to the rules you want to evaluate and then, click on the **evaluate** button. If you don't tick any rule, the evaluator will evaluate all the rules.

Delete or Refine selected rules (you can only refine underspecified rules that contain a regexp)

- none
- delete
- refine

save updates

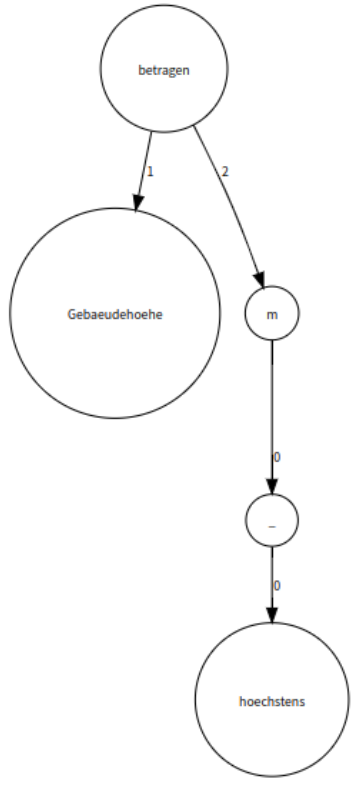
evaluate selected

Graph viewer and evaluator

Browse graphs: +

Choose from the rules

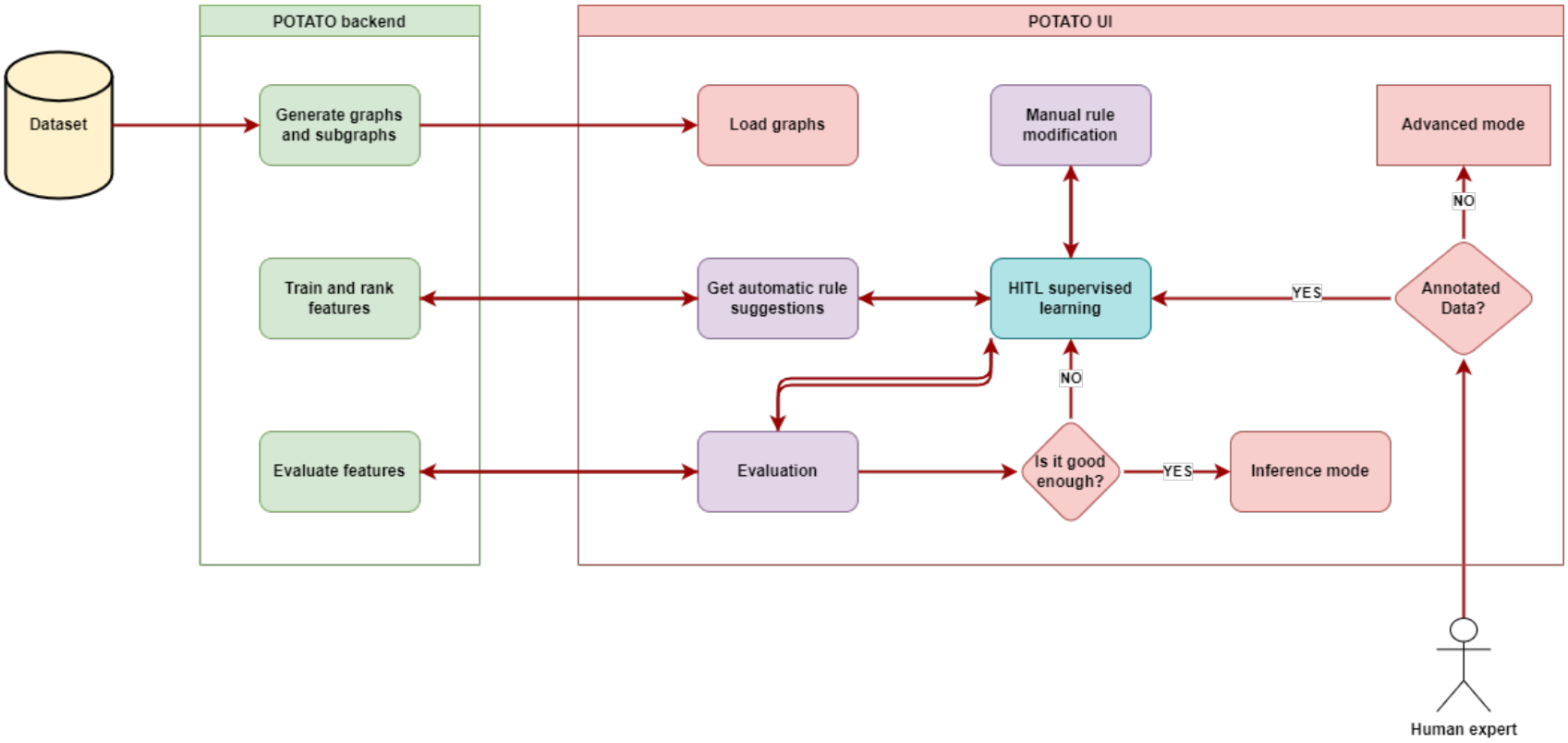
(u_482 / betragen :2 (u_96 / m :0 (u_744 / .* :0 (u_127 / hoechstens))) :1 (u_123 / Gebaeudehoehe))



Result of using all the rules: Precision: **0.975**, Recall: **0.765**, Fscore: **0.857**

The rule's result: Precision: **1.000**, Recall: **0.196**, Fscore: **0.328**, True positives: **10**, False positives: **0**

POTATO workflow



Thank you!

Questions?

Code:

- https://github.com/recski/tuw-nlp
- https://github.com/adaamko/POTATO

Reading:

- POTATO: paper [6] and blogpost [5]
- On building semantic graphs [12]

References I

[1] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21. Virtual Event, Canada: Association for Computing Machinery, 2021, pp. 610–623. ISBN: 9781450383097. DOI: 10.1145/3442188.3445922. URL: https://doi.org/10.1145/3442188.3445922.

[2] Sarthak Jain and Byron C. Wallace. "Attention is not Explanation". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 3543–3556. DOI: 10.18653/v1/N19-1357. URL: https://aclanthology.org/N19-1357.

[3] Robin Jia and Percy Liang. "Adversarial Examples for Evaluating Reading Comprehension Systems". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics, 2017, pp. 2021–2031. DOI: 10.18653/v1/D17-1215. URL: https://www.aclweb.org/anthology/D17-1215.

[4] Dan Jurafsky and James H Martin. "Speech and language processing". In: (2009).

[5] Ádám Kovács. *Using POTATO for interpretable information extraction*. https://towardsdatascience.com/using-potato-for-interpretable-information-extraction-f2081a717eb7. 2022.

[6] Ádám Kovács, Kinga Gémes, Eszter Iklódi, and Gábor Recski. *POTATO: exPlainable infOrmation exTrAcTion framewOrk*. 2022. DOI: 10.48550/ARXIV.2201.13230. URL: https://arxiv.org/abs/2201.13230.

[7] Keita Kurita, Nidhi Vyas, Ayush Pareek, Alan W Black, and Yulia Tsvetkov. "Measuring Bias in Contextualized Word Representations". In: *Proceedings of the First Workshop on Gender Bias in Natural Language Processing*. Florence, Italy: Association for Computational Linguistics, 2019, pp. 166–172. DOI: 10.18653/v1/W19-3823. URL: https://www.aclweb.org/anthology/W19-3823.

[8] Hung-yi Lee. *Recurrent Neural Network (RNN)*. http://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2016/Lecture/RNN%20(v2).pdf. 2016.

References II

[9] Aleksander Obuchowski. *New state-of-the-art intent detection model from SentiOne*. https://sentione.com/blog/new-state-of-the-art-intent-detection-model-from-sentione. 2021.

[10] Simon Ostermann, Michael Roth, Ashutosh Modi, Stefan Thater, and Manfred Pinkal. "SemEval-2018 Task 11: Machine Comprehension Using Commonsense Knowledge". In: *Proceedings of The 12th International Workshop on Semantic Evaluation*. New Orleans, Louisiana: Association for Computational Linguistics, 2018, pp. 747–757. DOI: 10.18653/v1/S18-1119. URL: https://aclanthology.org/S18-1119.

[11] Danish Pruthi, Mansi Gupta, Bhuwan Dhingra, Graham Neubig, and Zachary C. Lipton. "Learning to Deceive with Attention-Based Explanations". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, 2020, pp. 4782–4793. DOI: 10.18653/v1/2020.acl-main.432. URL: https://www.aclweb.org/anthology/2020.acl-main.432.

[12] Gábor Recski. "Building concept definitions from explanatory dictionaries". In: *International Journal of Lexicography* 31 (3 2018), pp. 274–311. DOI: 10.1093/ijl/ecx007. URL: https://academic.oup.com/ijl/article/31/3/274/3835852?guestAccessKey=9f090231-2795-47af-aa44-5fb15d4df0d8.

[13] Shachar Rosenman, Alon Jacovi, and Yoav Goldberg. "Exposing Shallow Heuristics of Relation Extraction Models with Challenge Data". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics, 2020, pp. 3702–3710. DOI: 10.18653/v1/2020.emnlp-main.302. URL: https://aclanthology.org/2020.emnlp-main.302.

[14] Aarne Talman and Stergios Chatzikyriakidis. "Testing the Generalization Power of Neural Network Models across NLI Benchmarks". In: *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*. Florence, Italy: Association for Computational Linguistics, Aug. 2019, pp. 85–94. DOI: 10.18653/v1/W19-4810. URL: https://www.aclweb.org/anthology/W19-4810.

References III

[15] Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, and Christopher D. Manning. "Position-aware Attention and Supervised Data Improve Slot Filling". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*. 2017, pp. 35–45. URL: https://nlp.stanford.edu/pubs/zhang2017tacred.pdf.

Mapping from syntax to semantics

Dependency	Edge
advcl, advmod, amod, ...	$w_1 \xrightarrow{0} w_2$
nsubj, csubj, ...	$w_1 \underset{0}{\overset{1}{\rightleftharpoons}} w_2$
obj, ccomp, ...	$w_1 \xrightarrow{2} w_2$
nmod:poss	$w_2 \xleftarrow{1} \texttt{HAS} \xrightarrow{2} w_1$
*:tmod	$w_1 \xleftarrow{1} \texttt{AT} \xrightarrow{2} w_2$
...	
$w_1 \xrightarrow{\text{obl}} w_2 \xrightarrow{\text{case}} w_3$	$w_1 \xleftarrow{1} w_3 \xrightarrow{2} w_2$
$w_1 \xrightarrow{\text{acl:relcl}} w_2 \xrightarrow{\text{nsubj}} w_3$	$w_1 \xrightarrow{0} w_2$