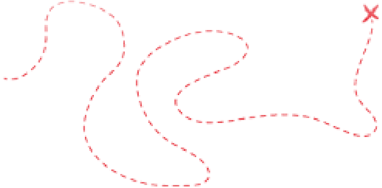
Meaning Representations and Where to Find Them

Daniel Hershcovich

ML Section
November 9, 2020

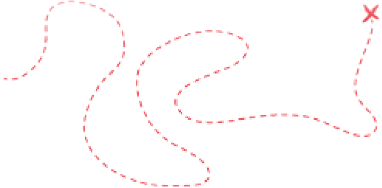
The Route

- Meaning representations are useful for NLP, and NLG evaluation.



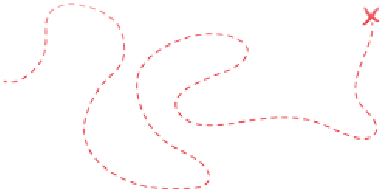
The Route

- Meaning representations are useful for NLP, and NLG evaluation.
- They have accurate parsers for various frameworks and languages.



The Route

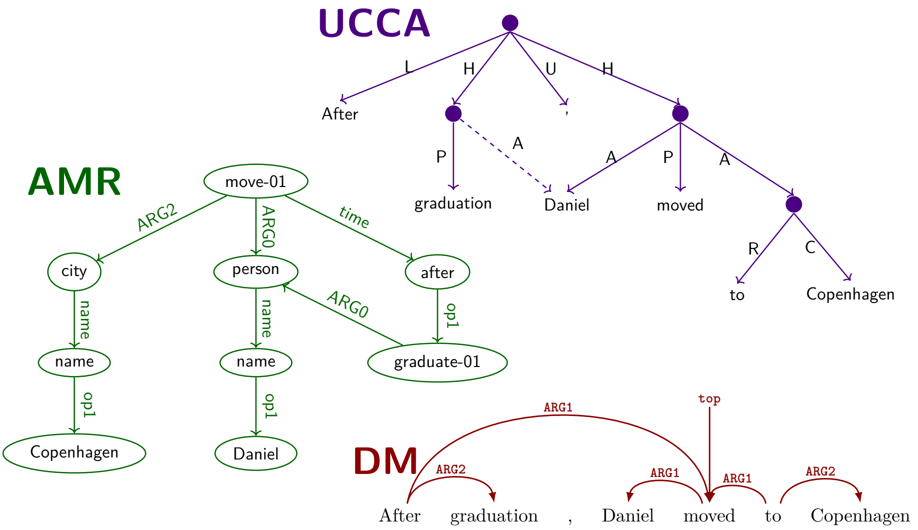
- Meaning representations are useful for NLP, and NLG evaluation.
- They have accurate parsers for various frameworks and languages.
- Linguistic and supervised conversion enable deep comparison.



Outline

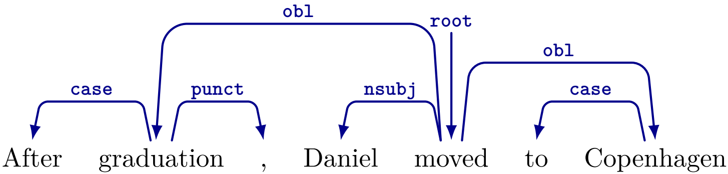
1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

Meaning Representations



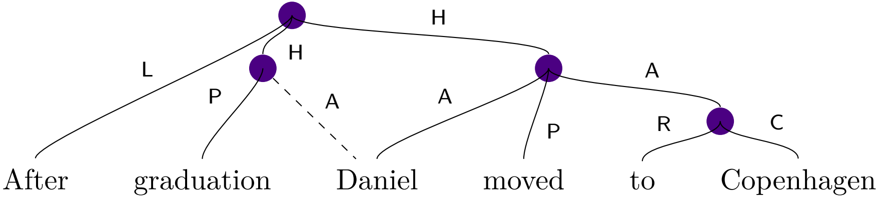
Syntactic Representations

UD (Universal Dependencies)



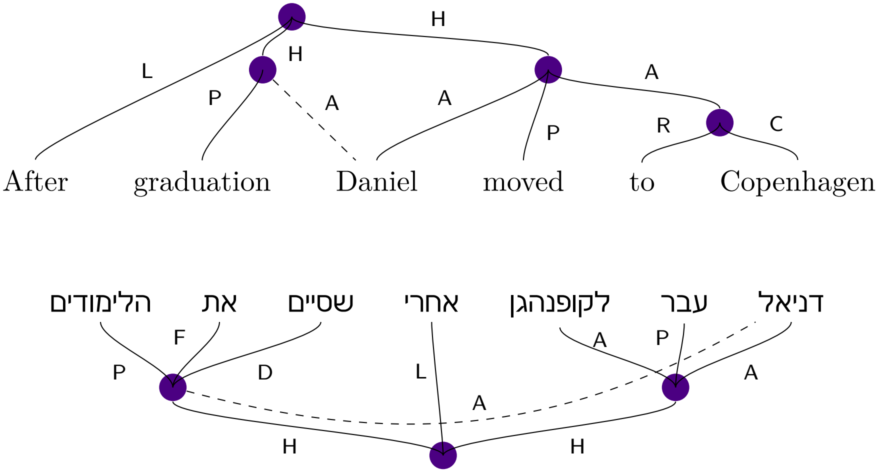
Universal Conceptual Cognitive Annotation (UCCA)

Supports rapid and intuitive annotation of linguistic semantic phenomena. [Abend and Rappoport, 2013]



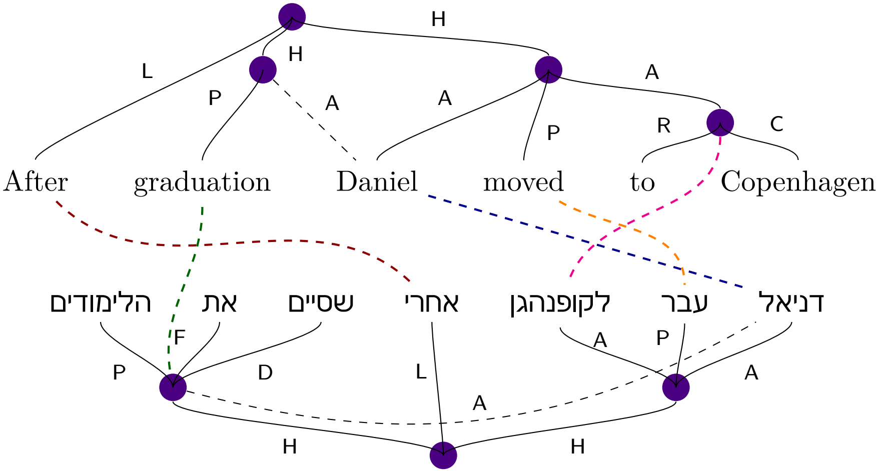
Universal Conceptual Cognitive Annotation (UCCA)

Supports rapid and intuitive annotation of linguistic semantic phenomena.
Cross-linguistically applicable and stable [Sulem et al., 2015].



Universal Conceptual Cognitive Annotation (UCCA)

Supports rapid and intuitive annotation of linguistic semantic phenomena.
Cross-linguistically applicable and stable [Sulem et al., 2015].



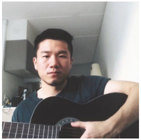
Universal Conceptual Cognitive Annotation (UCCA)

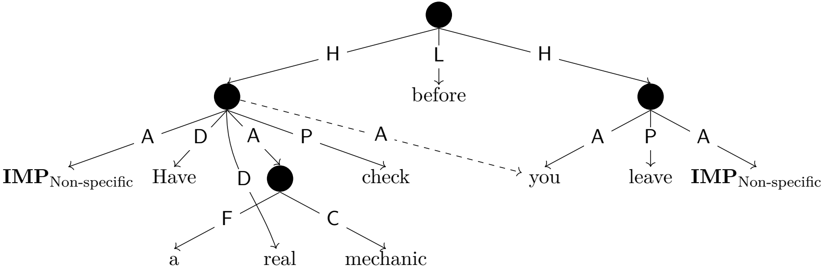
UCCA tutorial at COLING 2020:

`http://bit.ly/ucca-tutorial`

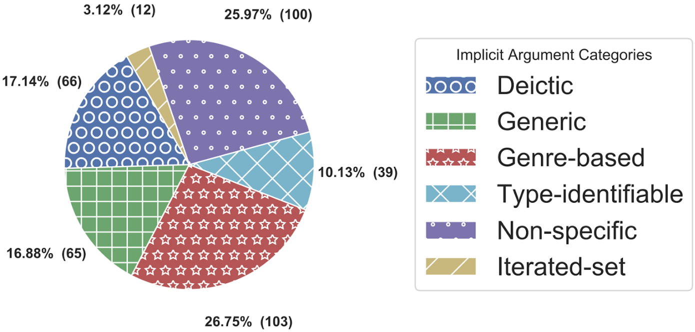
Refining Implicit Argument Annotation for UCCA

Fine-grained annotation of EWT corpus [Cui and Hershcovich, 2020].





Refining Implicit Argument Annotation for UCCA



What can meaning representation do for NLP?

- Incorporating linguistically informed rules
- Controlled evaluation by explicit criteria
- Inductive bias to facilitate learning
- Explainable models by design

Outline

Sentence Splitting for Text Simplification

Last year I read the book John authored →

John wrote a book. I read the book.

Sentence Splitting for Text Simplification

Last year I read the book John authored →

John wrote a book. I read the book.

MT-based simplifcation is *overly conservative*.

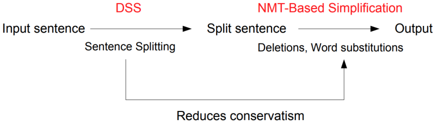
Sentence Splitting for Text Simplification

Last year I read the book John authored $\rightarrow$

John wrote a book. I read the book.

MT-based simplifcation is *overly conservative*.

Direct Semantic Splitting before MT-based simplification to place each scene in its own sentence [Sulem et al., 2018c].

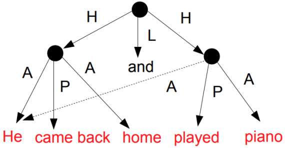


Sentence Splitting for Text Simplification

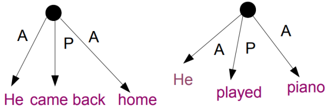
Rule 1:

The Semantic Rules

Parallel Scenes



He came back home and played piano.



He came back home. He played piano.

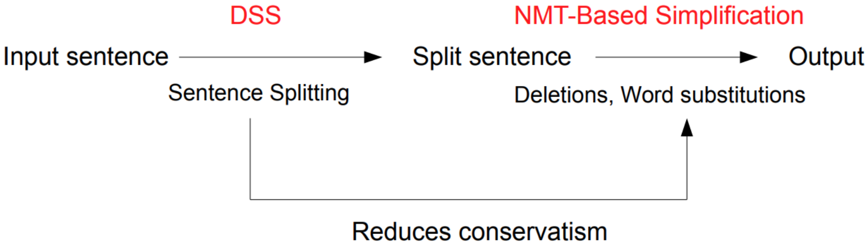
Sentence Splitting for Text Simplification

He observed the planet. Planet has 14 satellites.

Sentence Splitting for Text Simplification

He observed the planet. Planet has 14 satellites.

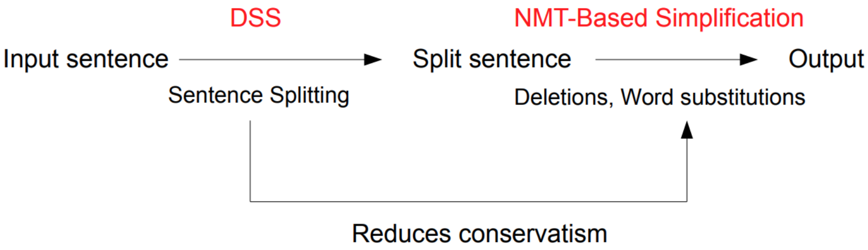
Neural MT methods to fix grammaticality.



Sentence Splitting for Text Simplification

He observed the planet. Planet has 14 satellites.

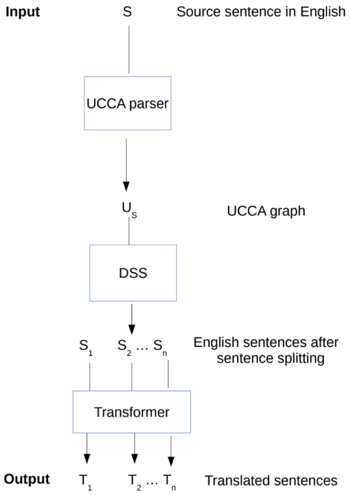
Neural MT methods to fix grammaticality.



He observed the planet. **The planet** has 14 satellites.

Semantic Structural Decomposition for NMT

Splitting before NMT improves fluency [Sulem et al., 2020].



Outline

1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

BLEU is Not Suitable for the Evaluation

BLEU: reference-based evaluation metric for MT, also widely used to evaluate text simplification.

BLEU is Not Suitable for the Evaluation

BLEU: reference-based evaluation metric for MT, also widely used to evaluate text simplification.

With sentence splitting, not correlated with grammaticality or meaning preservation [Sulem et al., 2018a].

BLEU is Not Suitable for the Evaluation

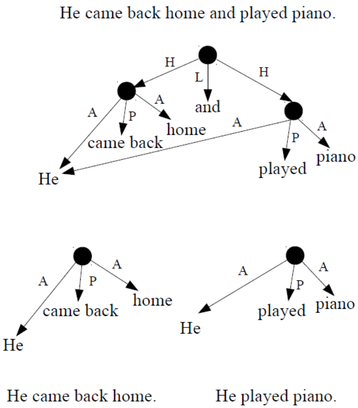
BLEU: reference-based evaluation metric for MT, also widely used to evaluate text simplification.

With sentence splitting, not correlated with grammaticality or meaning preservation [Sulem et al., 2018a].

Negatively correlated with simplicity!

Semantic Structural Evaluation for Text Simplification

SAMSA: reference-less measure of *structural simplicity* and *meaning preservation* [Sulem et al., 2018b].
Same principle: *one scene per sentence*.



Grammatical Error Correction

Another text-to-text generation task.

Ther is both sides of stories $\rightarrow$

There are two sides to every story

Grammatical Error Correction

Another text-to-text generation task.

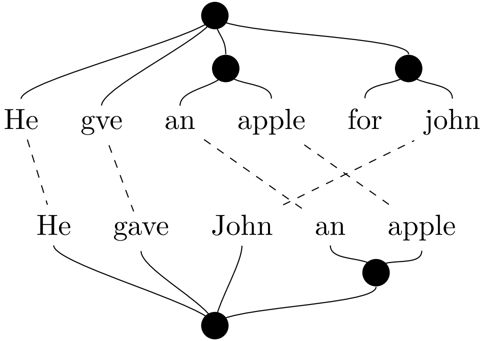
Ther is both sides of stories $\rightarrow$

There are two sides to every story

Using *references* for GEC evaluation encourages conservatism and underestimates precision [Choshen and Abend, 2018a].

Reference-less Measure of Faithfulness for GEC

- UCCA is *applicable* to ungrammatical learner language!
- UCCA is *stable* with respect to grammar corrections



Reference-less Measure of Faithfulness for GEC

USim measures meaning preservation automatically *without references* [Choshen and Abend, 2018b].

Reference-less Measure of Faithfulness for GEC

USim measures meaning preservation automatically *without references* [Choshen and Abend, 2018b].

Variation on standard UCCA evaluation, using unit *alignment* between the source and target graphs.

Reference-less Measure of Faithfulness for GEC

USim measures meaning preservation automatically *without references* [Choshen and Abend, 2018b].

Variation on standard UCCA evaluation, using unit *alignment* between the source and target graphs.

Sensitive to faithfulness, not overly conservative.

Source	the good student must know how to understand and work hard to get the iede.
Reference	A good student must be able to understand and work hard to get the idea.
Corrector	The good student must know how to understand and work hard to get on.

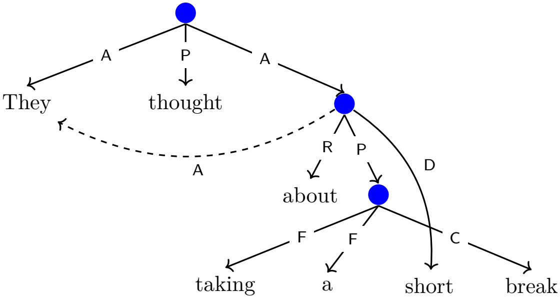
Outline

UCCA Parsing

The Task: Given plain text, predict its UCCA graph representation.

They thought about taking a short break

$\Downarrow$

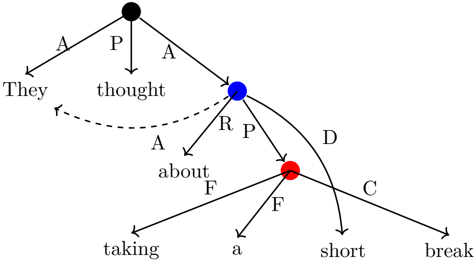


Outline

1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

TUPA

A Transition-Based Directed Acyclic Graph Parser for UCCA [Hershcovich et al., 2017].

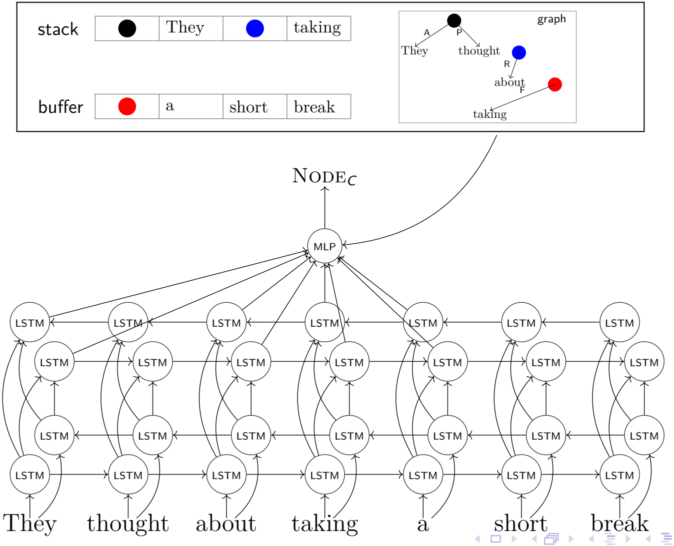


⇕

Shift, Right-Edge$_A$, Shift, Swap, Right-Edge$_P$, Reduce, Shift, Shift, Node$_R$, Reduce, Left-Remote$_A$, Shift, Shift, Node$_C$, Reduce, Shift, Right-Edge$_P$, Shift, Right-Edge$_F$, Reduce, Shift, Swap, Right-Edge$_D$, Reduce, Swap, Right-Edge$_A$, Reduce, Reduce, Shift, Reduce, Shift, Right-Edge$_C$, Finish

TUPA

Learns to greedily predict transition based on current state.



Data

UCCA training data is scarce

(English)

Data

UCCA training data is scarce

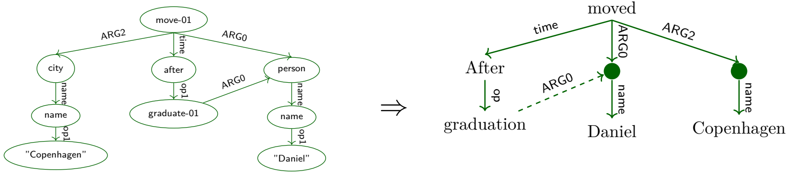
(English)

and domains are limited.

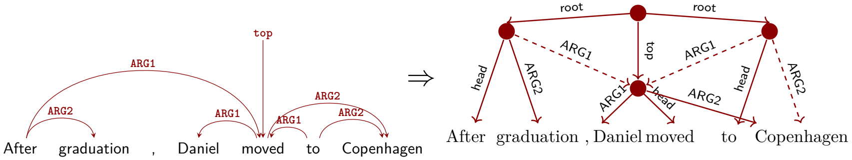
UCCA	AMR	DM	UD
Wikipedia	blogs	news	blogs
books	news		news
reviews	emails		emails
	reviews		reviews
			Q&A

Conversion

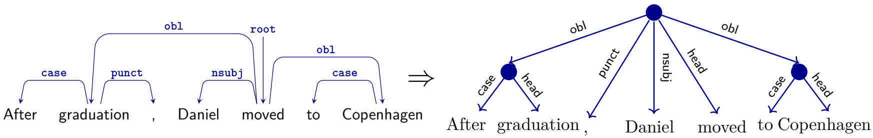
AMR



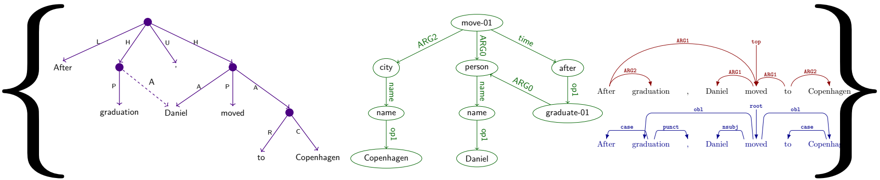
DM



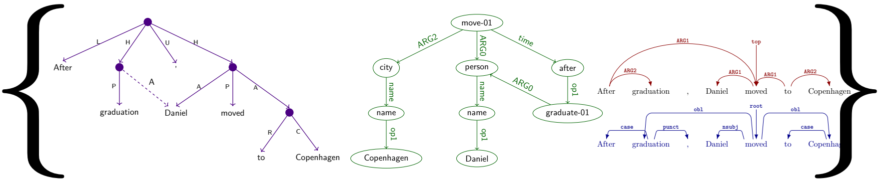
UD



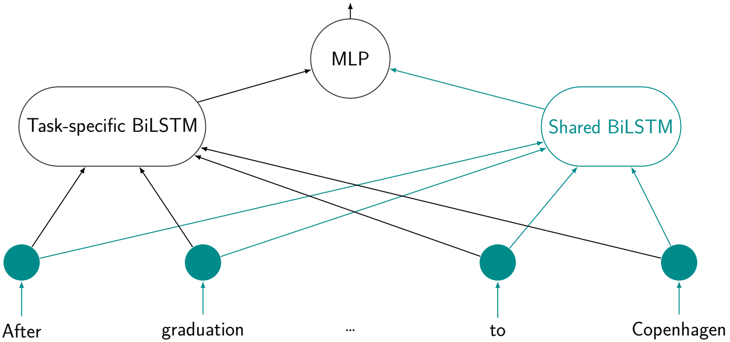
Multi-task


After
graduation
,
Daniel
moved
to
Copenhagen
L
H
U
H
P
A
A
P
A
R
C
move-01
ARG2
ARG0
time
city
person
after
name
name
op1
name
name
graduate-01
op1
op1
Copenhagen
Daniel
ARG1
top
ARG2
ARG1
ARG1
ARG2
obl
root
obl
case
punct
nsubj
case

Multi-task



Multi-task TUPA model [Hershcovich, Abend, and Rappoport, 2018]



Outline

1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

Shared tasks: parsing competitions

SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA [Hershcovich et al., 2019b]

- UCCA parsing in English, French and German.

Shared tasks: parsing competitions

SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA [Hershcovich et al., 2019b]

- UCCA parsing in English, French and German.

MRP 2019: Cross-Framework Meaning Representation Parsing [Oepen et al., 2019]

- DM, PSD, EDS, UCCA and AMR parsing in English.

Shared tasks: parsing competitions

SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA [Hershcovich et al., 2019b]

- UCCA parsing in English, French and German.

MRP 2019: Cross-Framework Meaning Representation Parsing [Oepen et al., 2019]

- DM, PSD, EDS, UCCA and AMR parsing in English.

MRP 2020: Cross-Framework and Cross-Lingual MRP [Oepen et al., 2020]

- EDS, PTG, UCCA, AMR and DRG parsing in English, Czech, German and Chinese.

SemEval 2019 Task 1: Cross-lingual Semantic Parsing

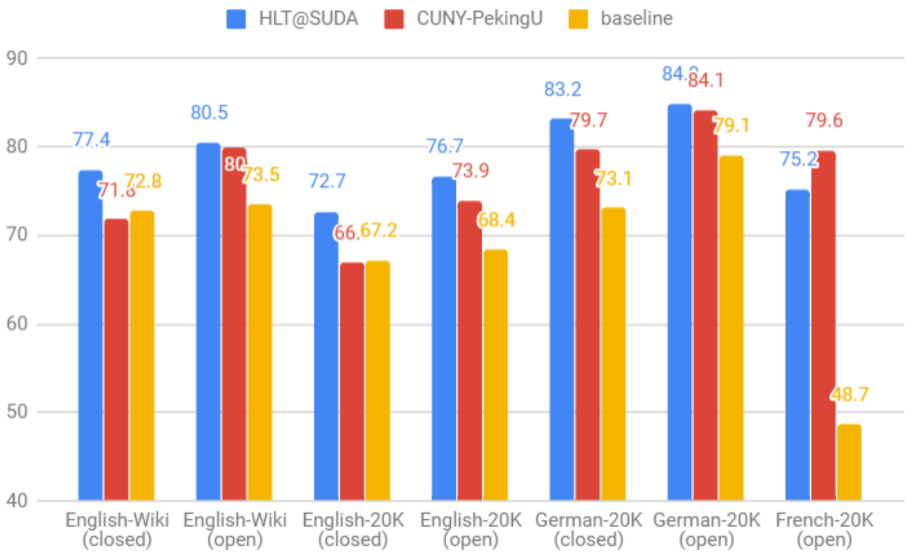
- UCCA parsing in English, French and German.
- 8 teams participated.
- Baseline: TUPA.

SemEval 2019 Task 1: Cross-lingual Semantic Parsing

- UCCA parsing in English, French and German.
- 8 teams participated.
- Baseline: TUPA.

- English {in-domain/out-of-domain} × {open/closed}
- German in-domain {open/closed}
- French *low-resource* (only 15 training sentences)

SemEval 2019 Task 1

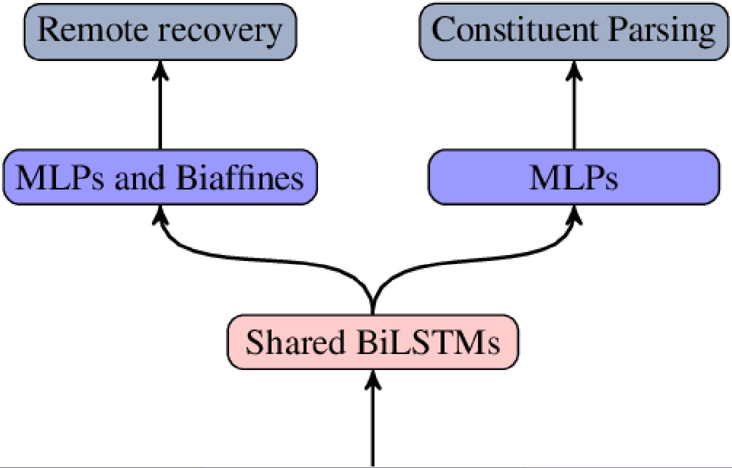


SemEval 2019 Task 1

HLT@SUDA:

Neural constituency parser + multi-task + BERT

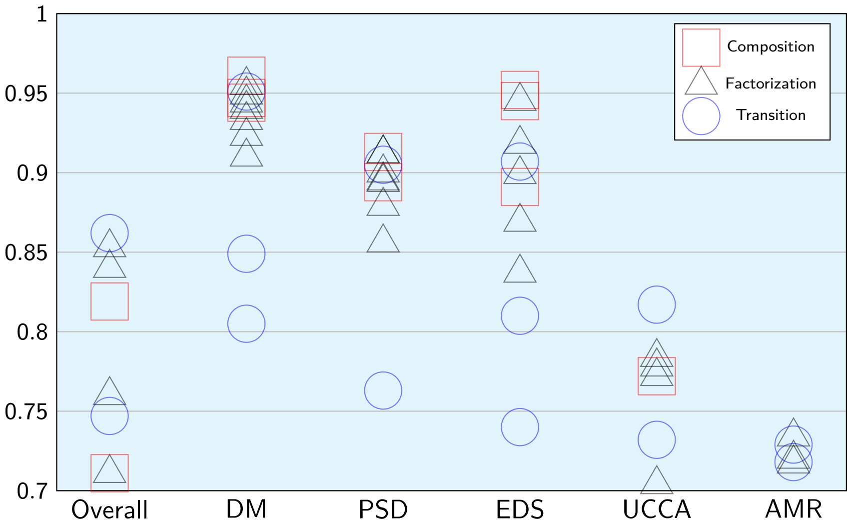
French: trained on all languages, with language embedding



MRP 2019

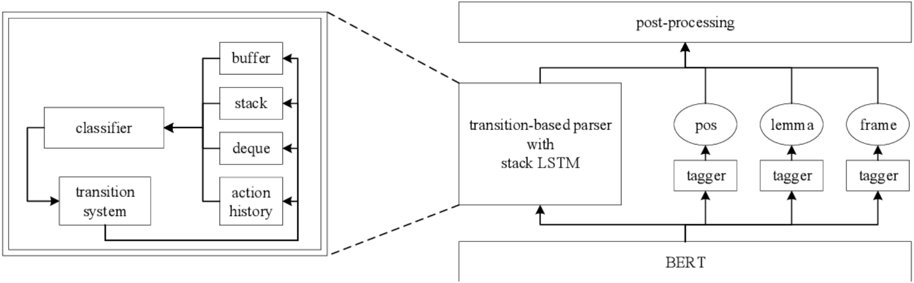
- DM, PSD, EDS, UCCA and AMR parsing in English.
- 18 teams participated.
- Baseline: TUPA (generalized beyond UCCA).

Results


1
0.95
0.9
0.85
0.8
0.75
0.7
Composition
Factorization
Transition
Overall
DM
PSD
EDS
UCCA
AMR

MRP 2019

Winning system: HIT-SCIR [Che et al., 2019].
Transition-based parser (similar to TUPA) + efficient training + BERT.



MRP 2020

- EDS, PTG, UCCA, AMR and DRG parsing in English, Czech, German and Chinese.
- 8 teams participated.

MRP 2020

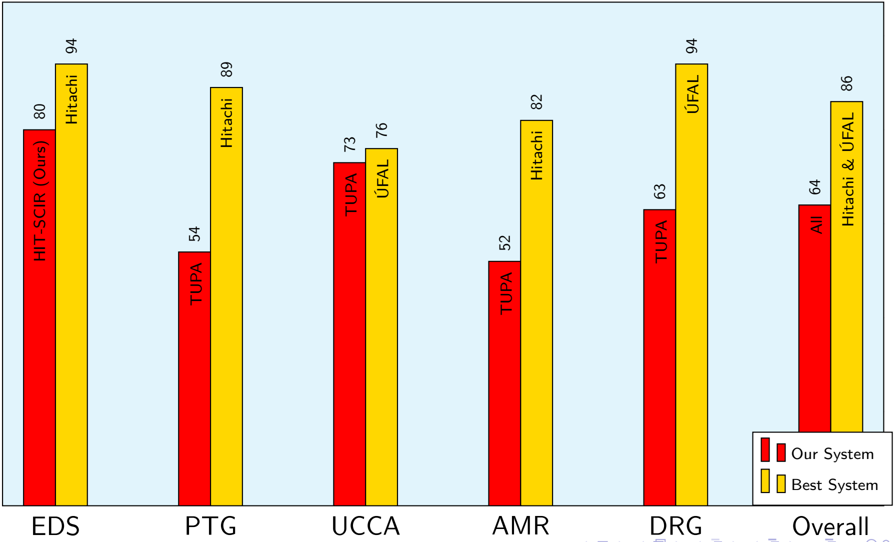
State of the art:

	EDS			UCCA			AMR		
	P	R	F	P	R	F	P	R	F
2019	.92	.93	.93	.84	.82	.83	.74	.72	.73
2020	.97	.97	.97	.86	.80	.83	.78	.79	.79

Winning systems: ÚFAL [Samuel and Straka, 2020] and Hitachi [Ozaki et al., 2020], encoder-decoder with pre-trained transformers.

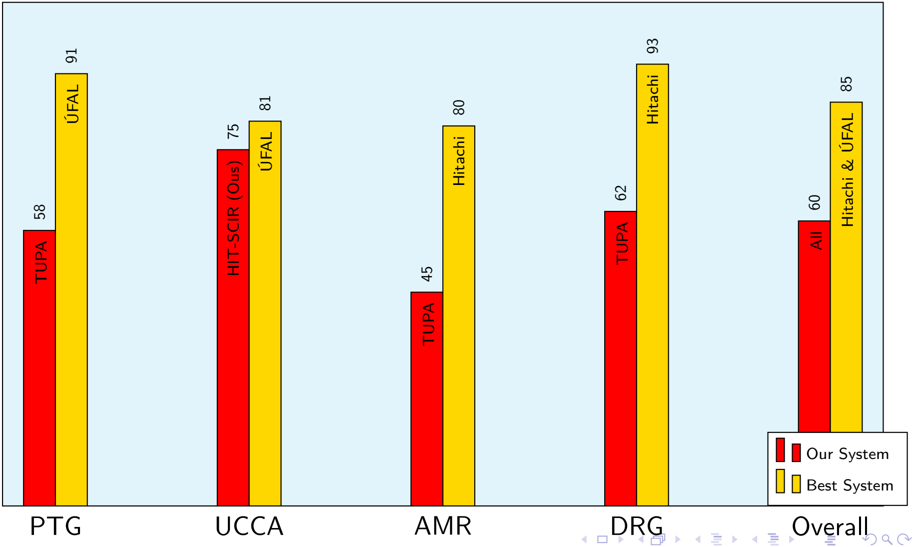
HUJI-KU at MRP 2020

Cross Framework Track:

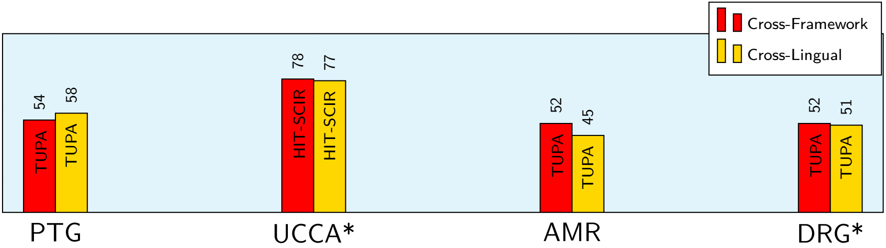


HUJI-KU at MRP 2020

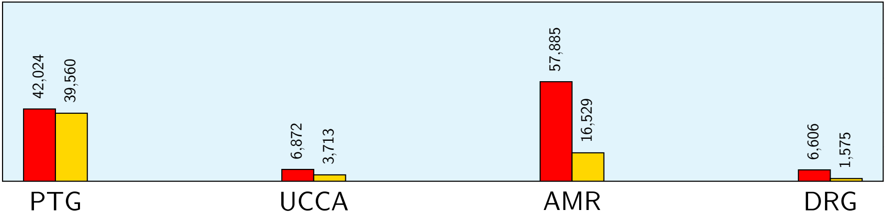
Cross Lingual Track:



CL vs. CF Track



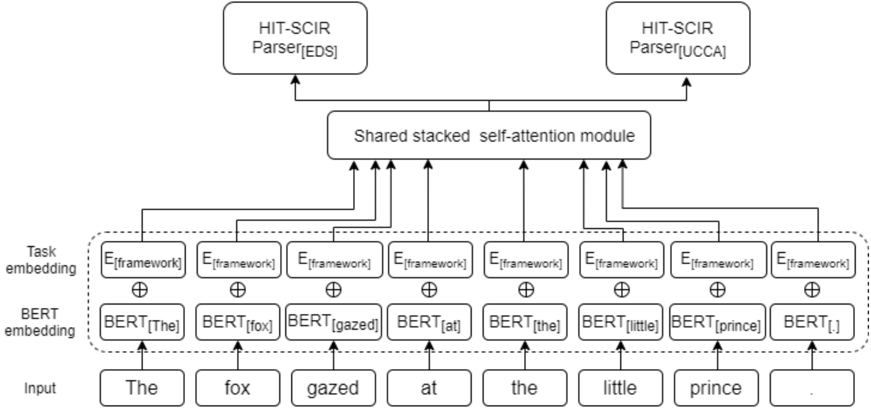
Full Evaluation (* for Validation) MRP F-score (%)



Dataset Size (# Sentences)

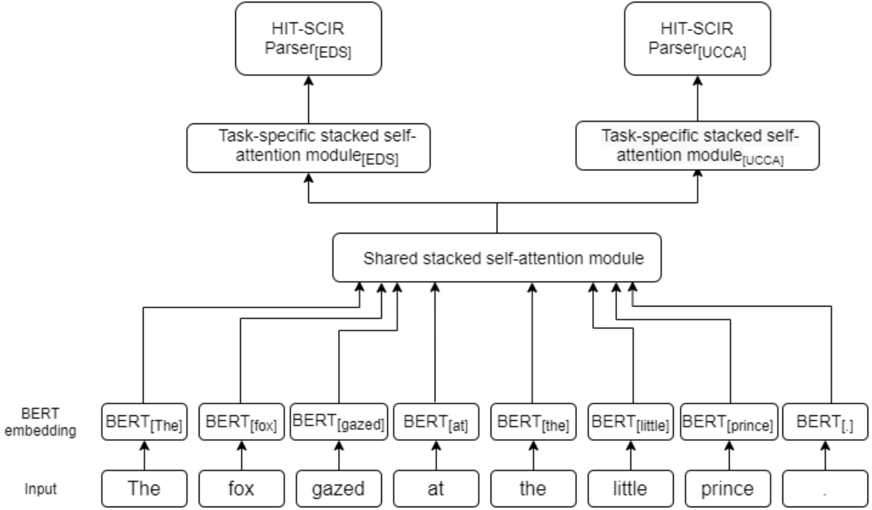
Multi-Task Model

Variant 1

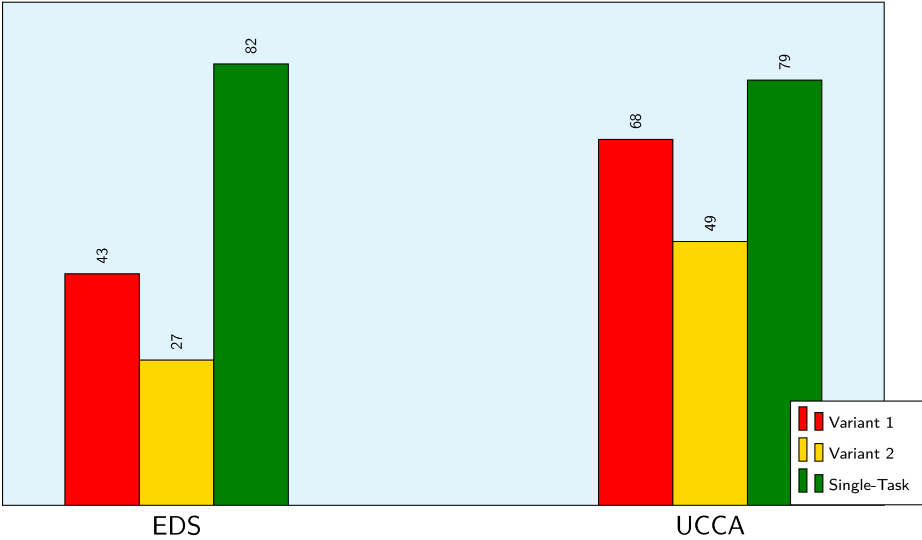


Multi-Task Model

Variant 2



Multi-Task Results



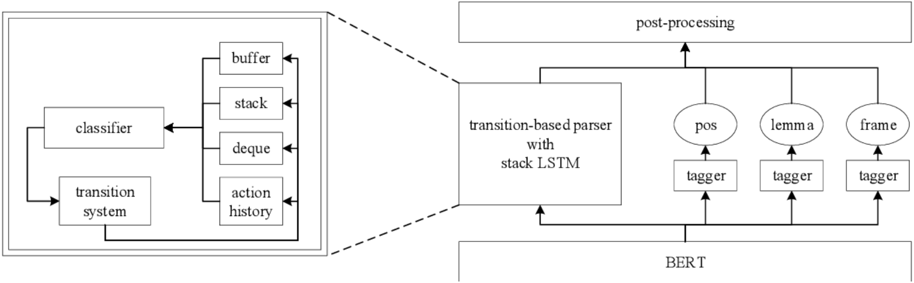
Cross-Framework Track Validation MRP F-score (%)

IWPT 2020

HIT-SCIR parser for **Enhanced UD** parsing [Hershcovich, de Lhoneux, Kulmizev, Pejhan, and Nivre, 2020a].

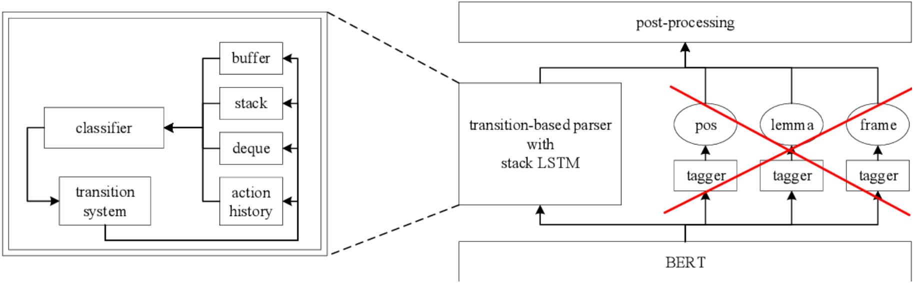
IWPT 2020

HIT-SCIR parser for **Enhanced UD** parsing [Hershcovich, de Lhoneux, Kulmizev, Pejhan, and Nivre, 2020a].

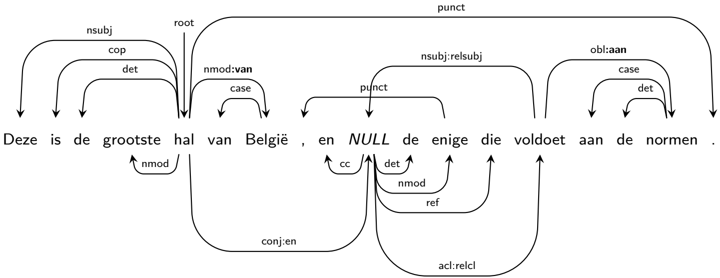


IWPT 2020

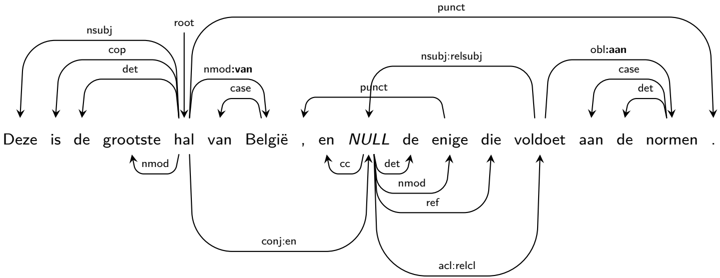
HIT-SCIR parser for **Enhanced UD** parsing [Hershcovich, de Lhoneux, Kulmizev, Pejhan, and Nivre, 2020a].



IWPT 2020

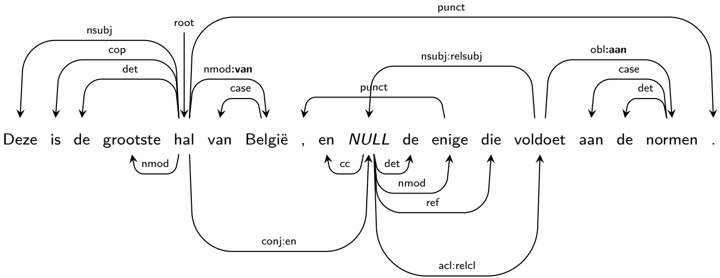

punct
root
nsubj
cop
det
nmod:**van**
case
nsubj:relsubj
punct
obl:**aan**
case
det
Deze is de grootste hal van België , en *NULL* de enige die voldoet aan de normen .
nmod
cc
det
nmod
ref
conj:en
acl:relcl

IWPT 2020



[ROOT] [Deze is de grootste hal van België , en de enige die voldoet aan de normen .]

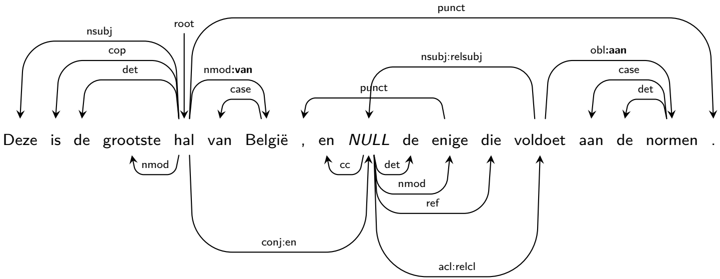
IWPT 2020



[ROOT] [Deze is de grootste hal van België , en de (...)]

IWPT 2020

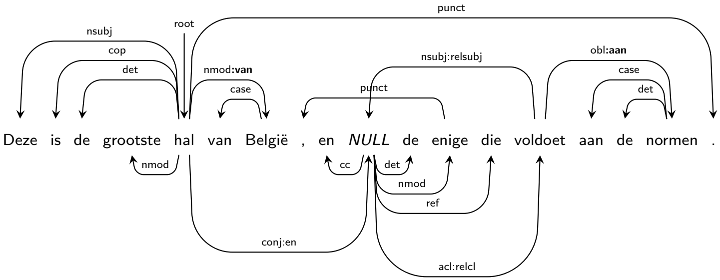
Shift



[ROOT Deze] [is de grootste hal van België , en de (...)]

IWPT 2020

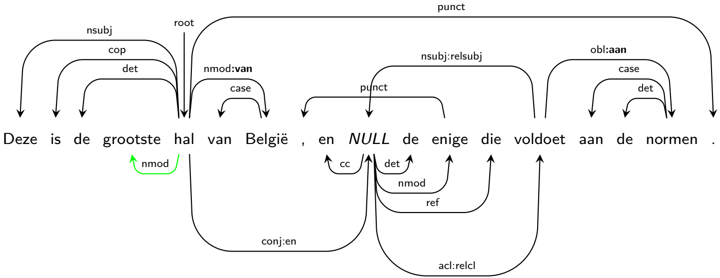
SHIFT x4



[ROOT Deze is de grootste hal] [van België , en de (...)]

IWPT 2020

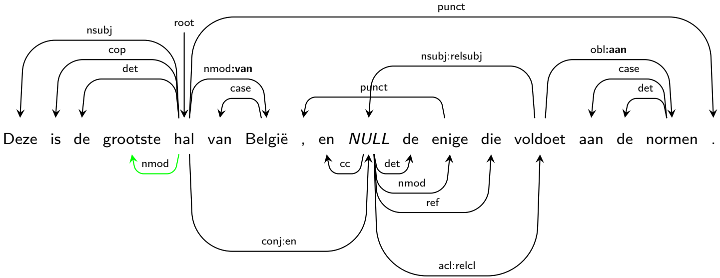
LEFT-EDGE:NMOD



[ROOT Deze is de grootste hal] [van België , en de (...)]

IWPT 2020

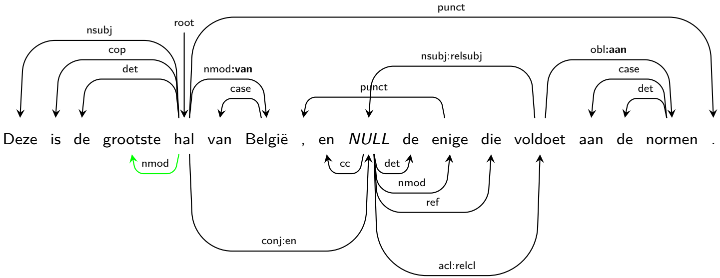
Node



[ROOT Deze is de grootste hal] [NULL van België , en (...)]

IWPT 2020

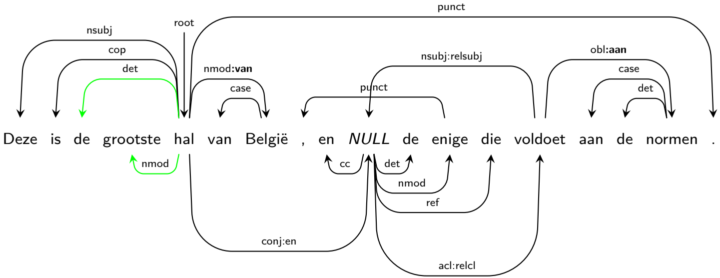
Reduce-1



[ROOT Deze is de hal] [NULL van België , en de enige (...)]

IWPT 2020

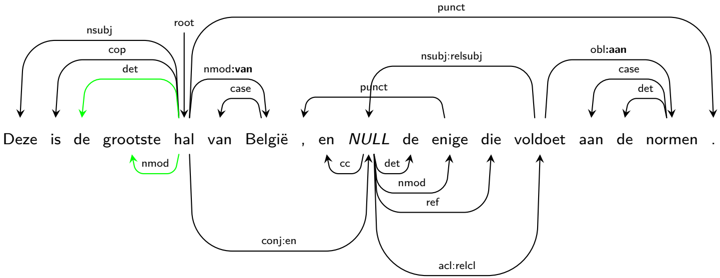
LEFT-EDGE:DET



[ROOT Deze is de hal] [NULL van België , en de enige (...)]

IWPT 2020

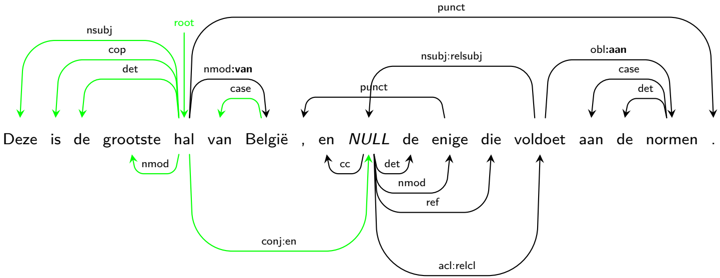
Reduce-1



[ROOT Deze is hal] [NULL van België , en de enige die (...)]

IWPT 2020

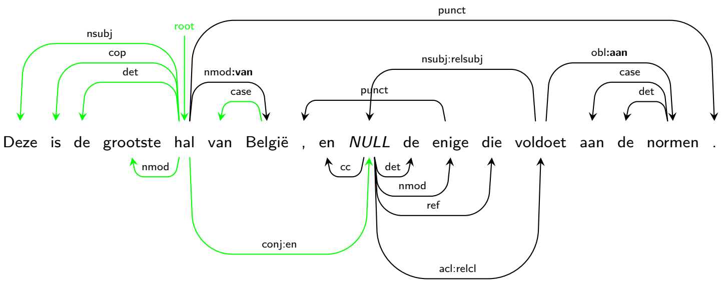
...



[ROOT hal NULL België] [, en de enige die voldoet aan (...)]

IWPT 2020

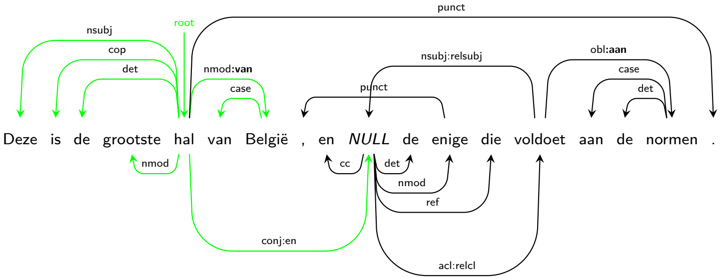
Swap



[ROOT hal België] [NULL , en de enige die voldoet aan (...)]

IWPT 2020

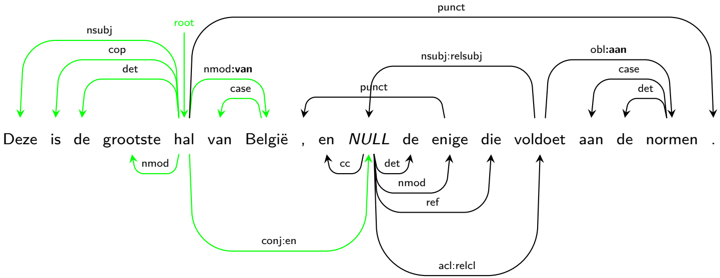
RIGHT-EDGE:NMOD:VAN



[ROOT hal België] [NULL , en de enige die voldoet aan (...)]

IWPT 2020

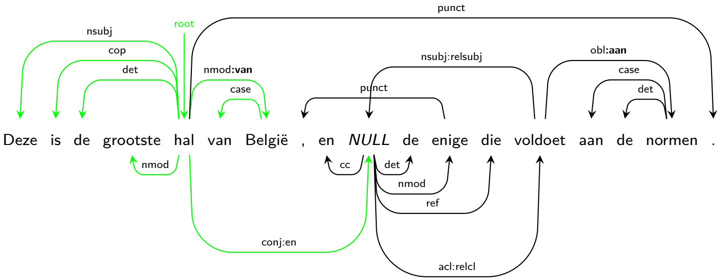
Reduce-0



[ROOT hal] [NULL , en de enige die voldoet aan de normen]

IWPT 2020

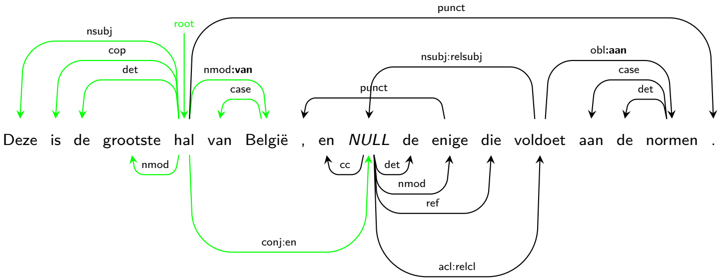
SHIFT x3



[ROOT hal NULL , en] [de enige die voldoet aan de normen]

IWPT 2020

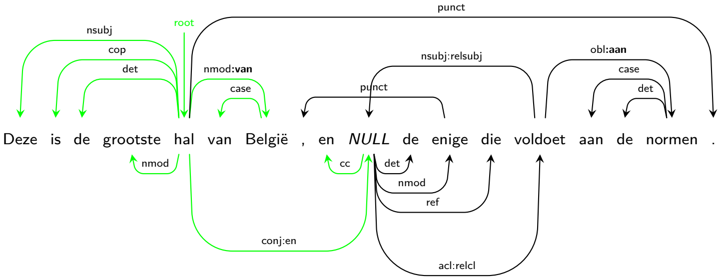
Swap



[ROOT hal NULL en] [, de enige die voldoet aan de normen]

IWPT 2020

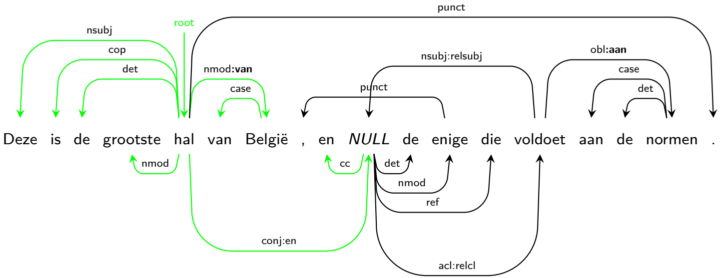
Right-Edge:cc



[ROOT hal NULL en] [, de enige die voldoet aan de normen]

IWPT 2020

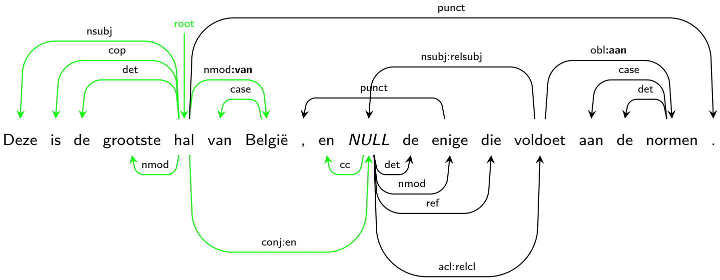
Reduce-0



[ROOT hal NULL] [, de enige die voldoet aan de normen]

IWPT 2020

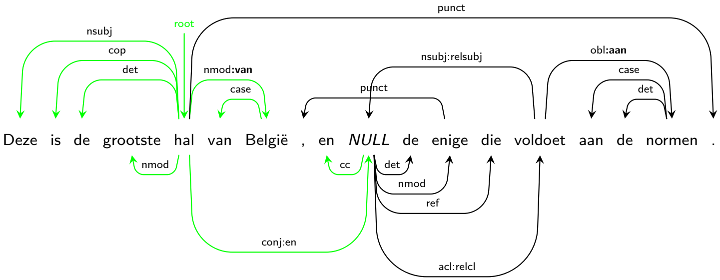
Shift x2



[ROOT hal NULL , de] [enige die voldoet aan de normen]

IWPT 2020

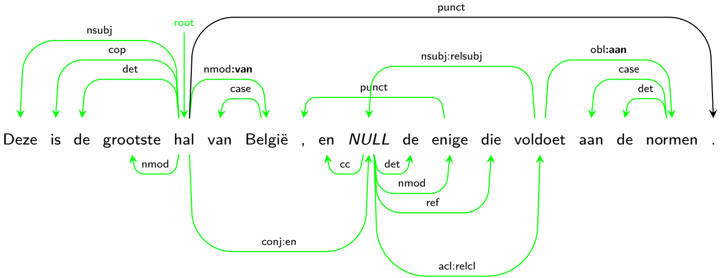
Swap



[ROOT hal NULL de] [, enige die voldoet aan de normen]

IWPT 2020

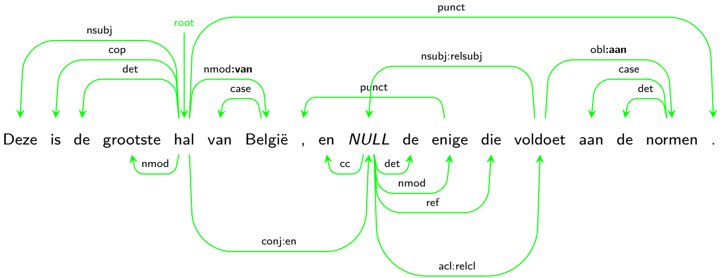
...



[ROOT hal .] [ ]

IWPT 2020

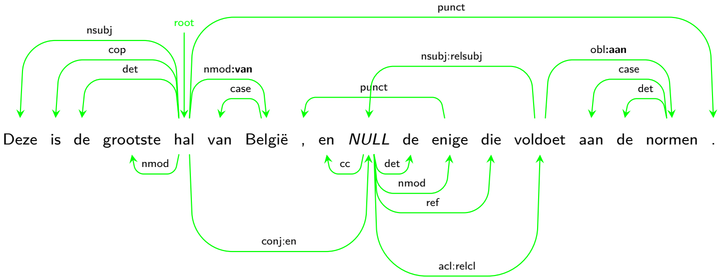
Right-Edge:punct



[ROOT hal .] [ ]

IWPT 2020

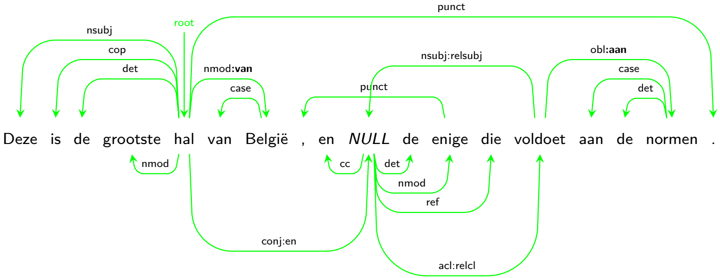
Reduce-0 x2



[ROOT] [ ]

IWPT 2020

Finish



[ROOT] [ ]

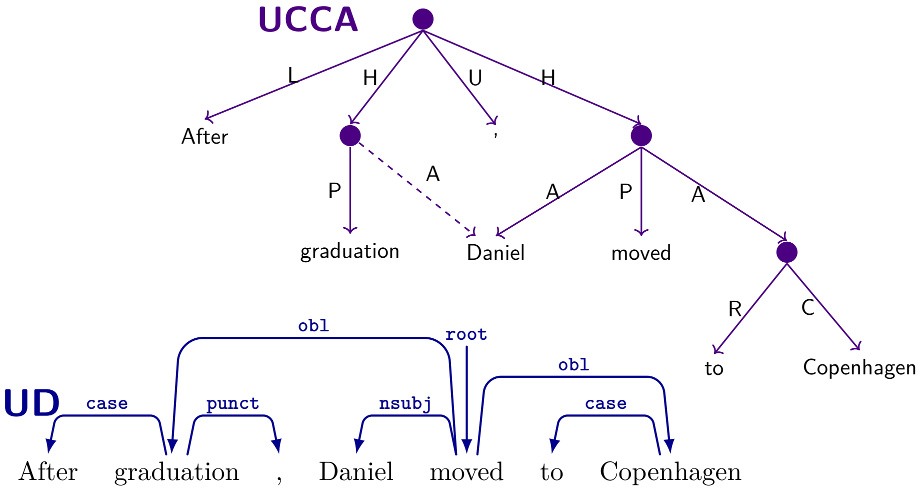
Outline

1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

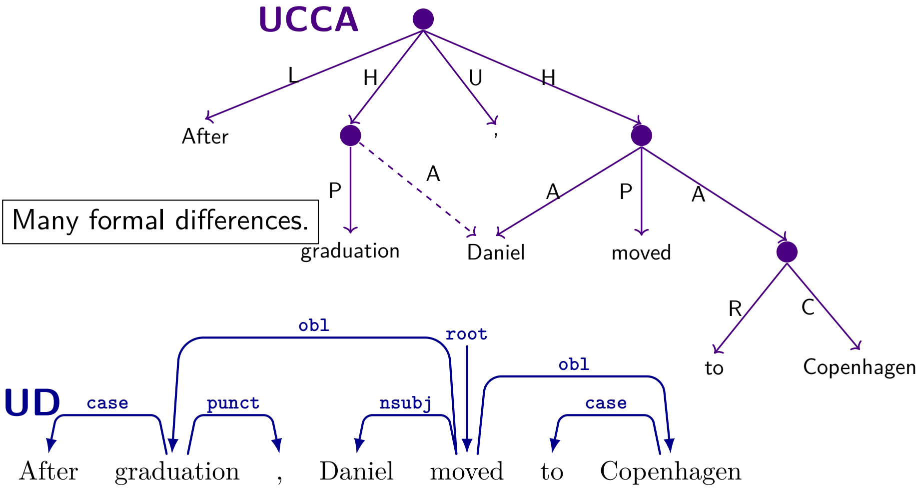
Outline

1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

UCCA vs. UD


UCCA
L
H
U
H
After
,
P
A
graduation
A
P
A
Daniel
moved
R
C
to
Copenhagen
UD
obl
root
obl
case
punct
nsubj
case
After graduation , Daniel moved to Copenhagen

UCCA vs. UD

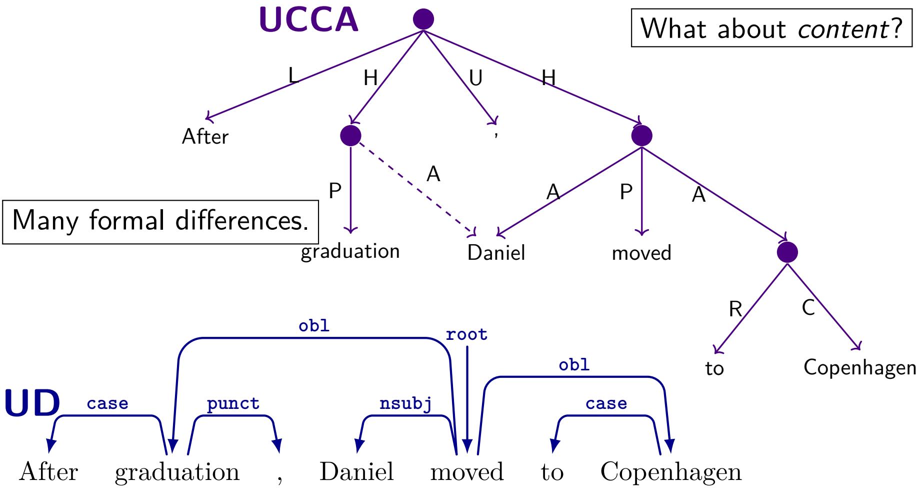


Many formal differences.

UCCA vs. UD

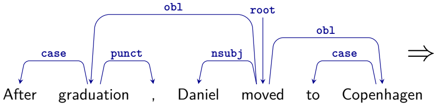
What about *content*?

Many formal differences.

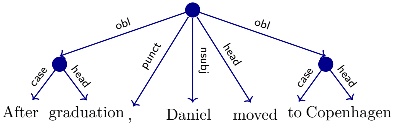


Assimilating the Graph Structures

UD

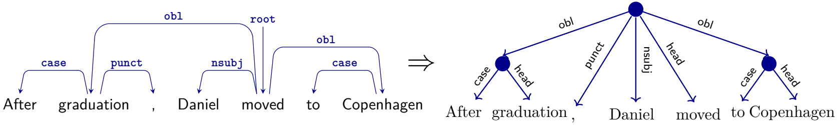


⇒



Assimilating the Graph Structures

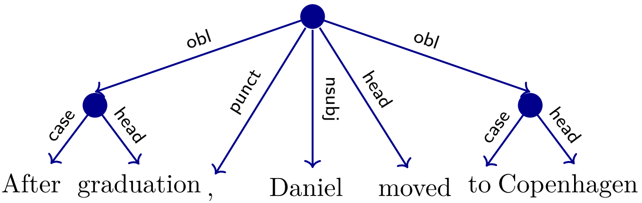
UD



Evaluate by matching edges [Hershcovich, Abend, and Rappoport, 2019a].

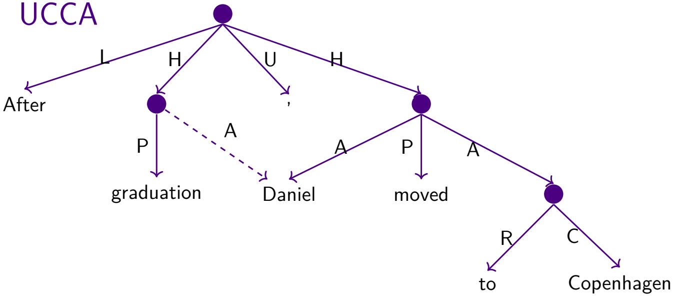
Assimilating the Graph Structures

UD



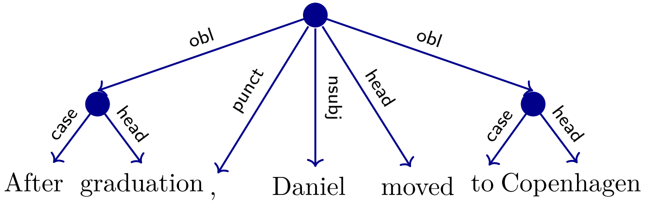
Evaluate by matching edges [Hershcovich, Abend, and Rappoport, 2019a].

UCCA



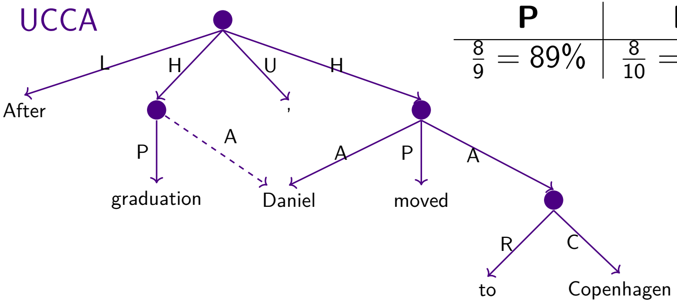
Assimilating the Graph Structures

UD



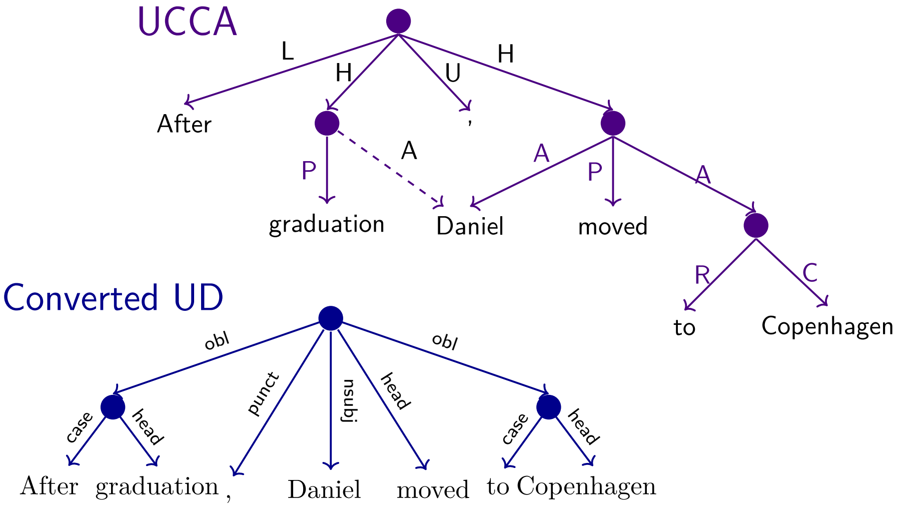
Evaluate by matching edges [Hershcovich, Abend, and Rappoport, 2019a].

UCCA

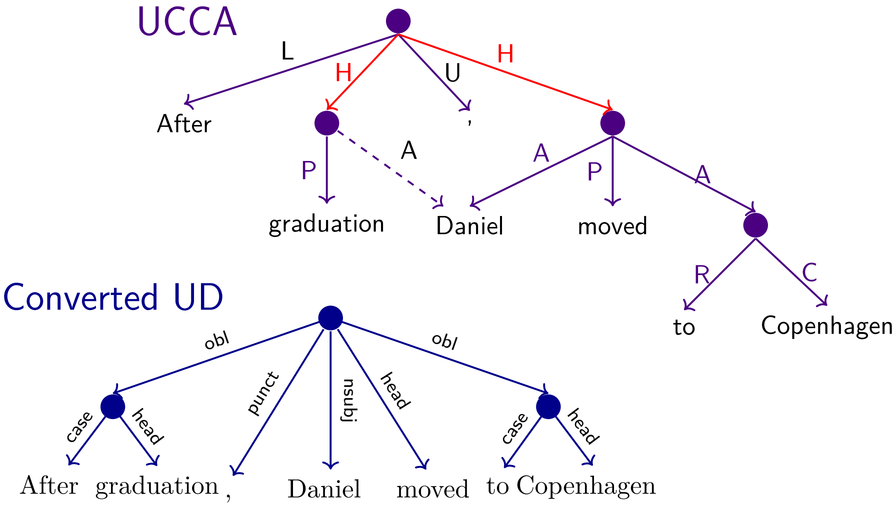


P	**R**	**F1**
$\frac{8}{9} = 89\%$	$\frac{8}{10} = 80\%$	84%

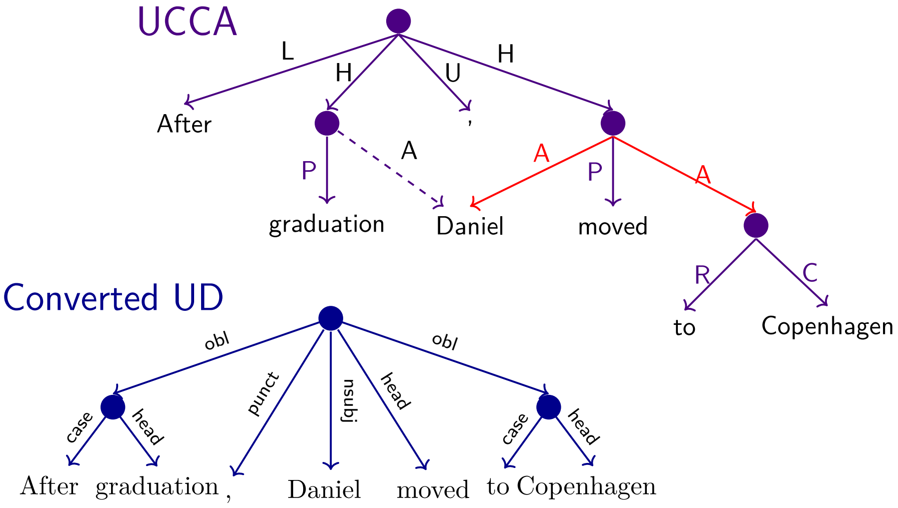
Scenes and non-Scenes, Relations and Participants



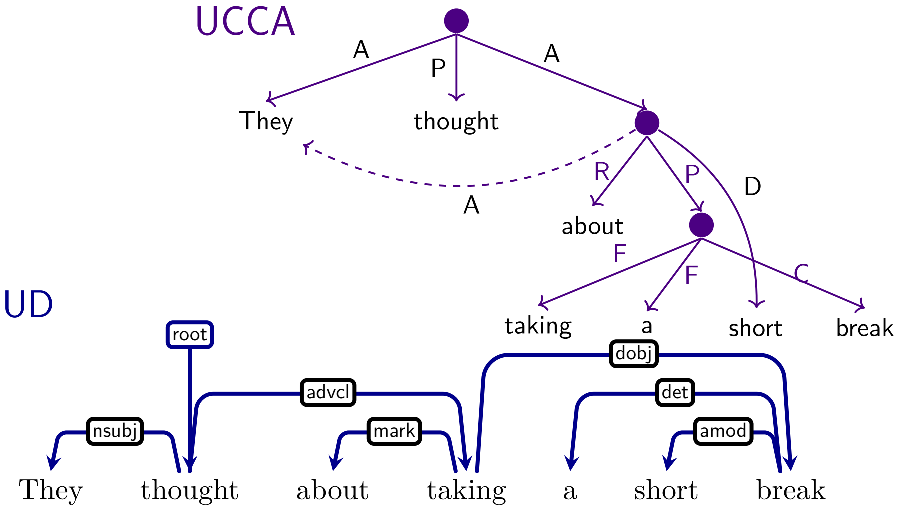
Scenes and non-Scenes, Relations and Participants



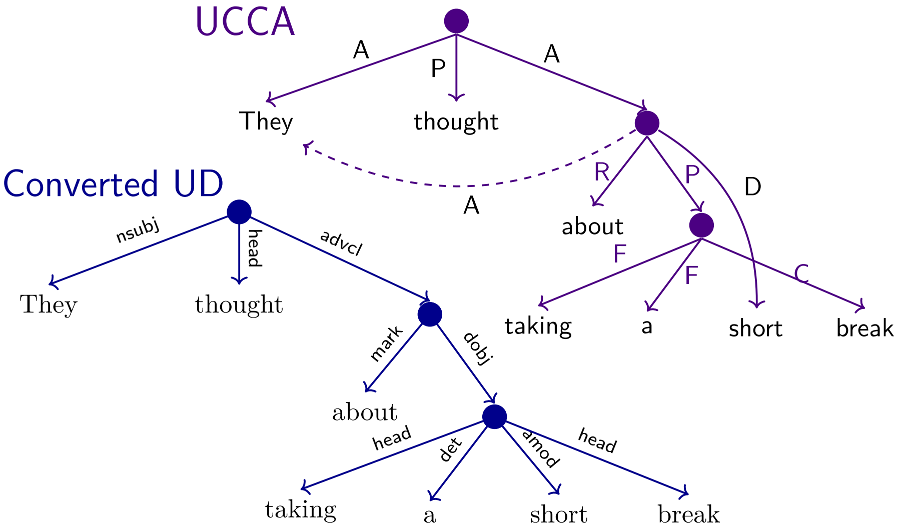
Scenes and non-Scenes, Relations and Participants



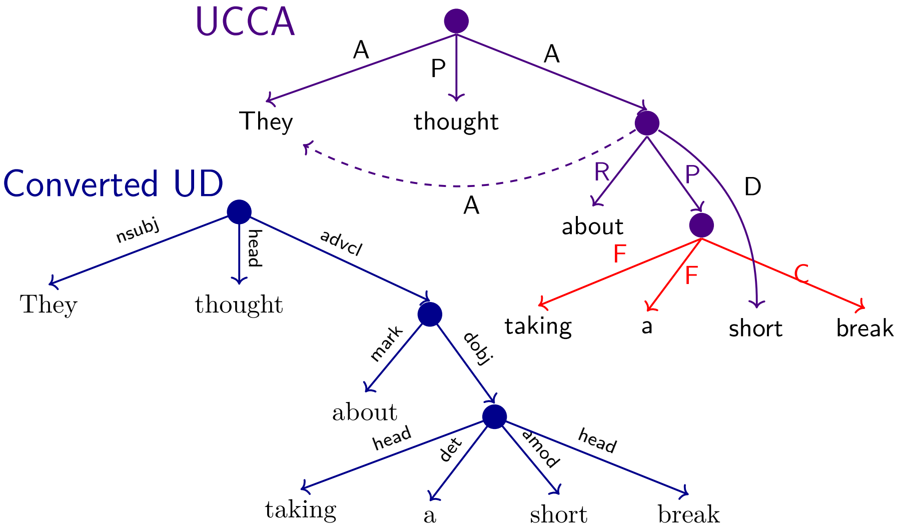
Multi-word Expressions



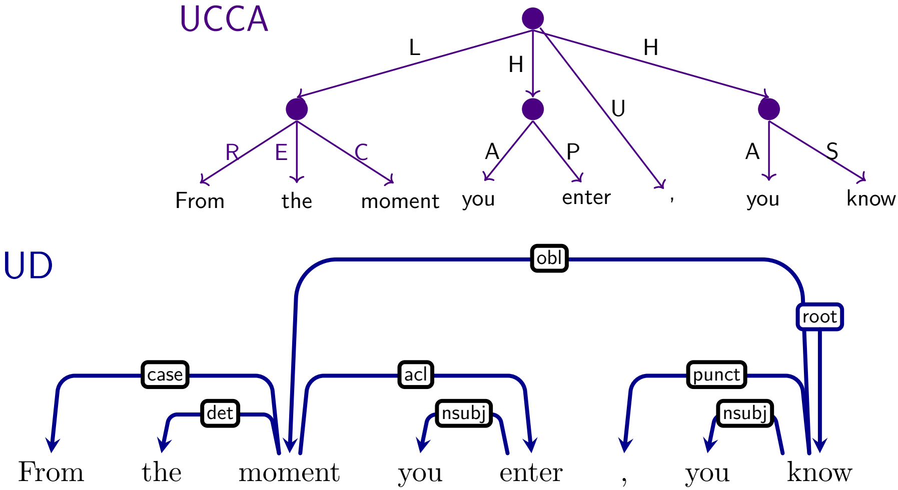
Multi-word Expressions



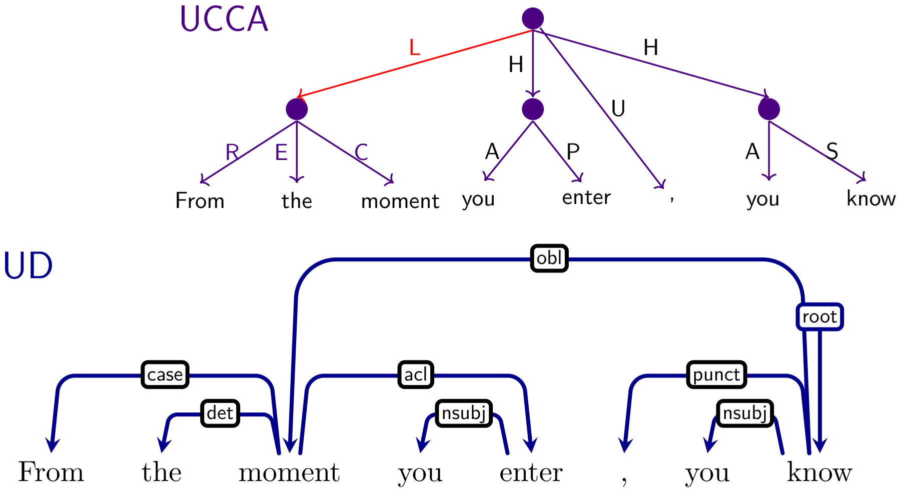
Multi-word Expressions



Linkage between Scenes



Linkage between Scenes

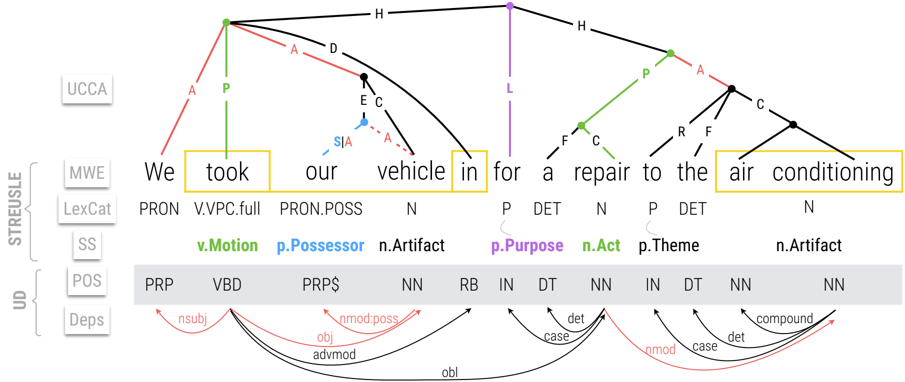


Outline

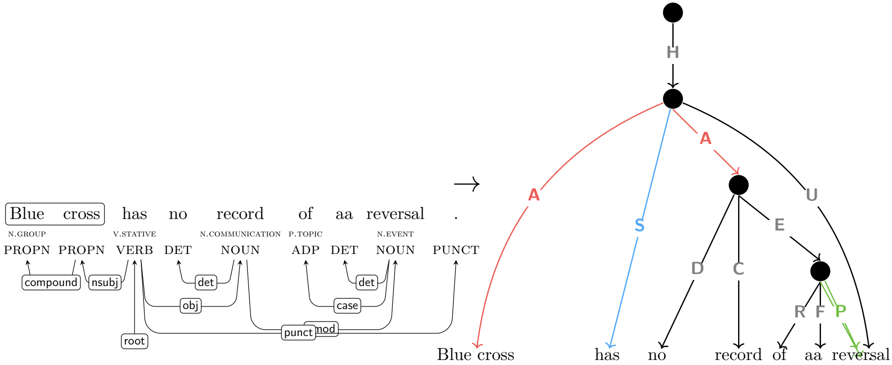
1. Representations
   - Incorporating linguistically informed rules into NLP
   - Controlled NLG evaluation by explicit criteria
2. Parsing
   - TUPA
   - Shared Tasks
3. Comparison
   - To Syntax
   - To Lexical Semantics

Comparison by Conversion: Reverse-Engineering UCCA from Syntax and Lexical Semantics

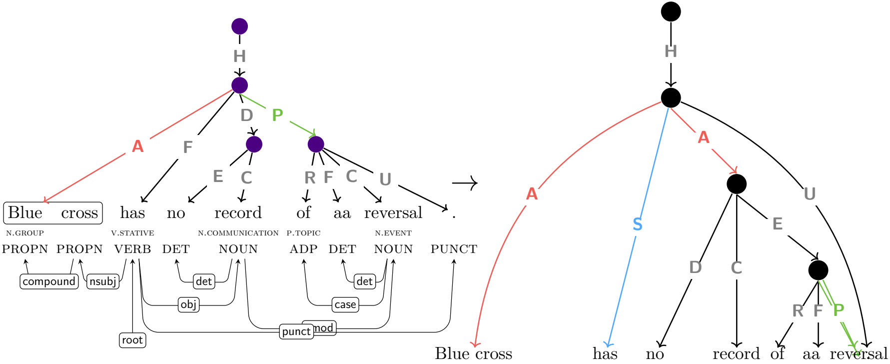
Complement syntax with *lexical* semantics to make up for differences [Hershcovich et al., 2020b].



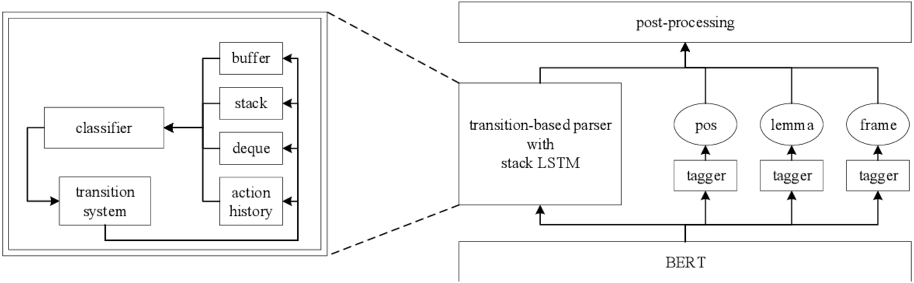
Comparison by Conversion



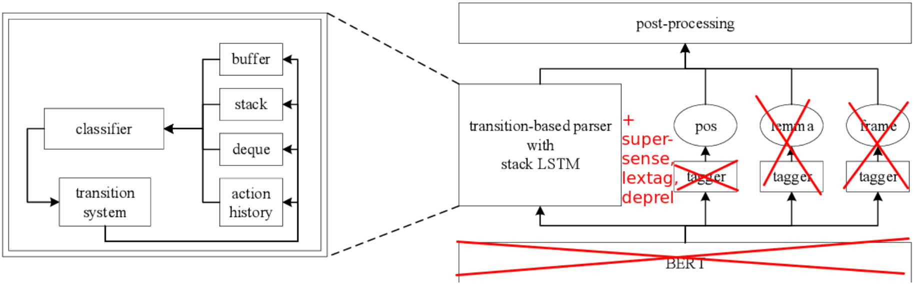
Comparison by Conversion



Comparison by Delexicalized Parsing



Comparison by Delexicalized Parsing



Comparison

	Primary F1	Remote F1
Syntax-based converter with UD	56.6	28.0
Our rule-based converter with UD+STREUSLE	71.7	44.2
TUPA, delex gold UD+STREUSLE	69.5	46.4
UD only	64.4	35.9
STREUSLE only	62.4	27.5
HIT-SCIR, delex with UD+STREUSLE	67.9	41.6
TUPA with UD + GloVe	71.7	47.0
HIT-SCIR (BERT-Large)	71.9	41.8
HIT-SCIR (GloVe)	67.0	42.4
with UD+STREUSLE	72.2	46.9

Confusion Matrix

Predicted Category \ Gold Category	A	A|G	A|P	A|S	C	D	D|T	E	F	G	H	L	N	P	Q	R	S	T	∅
A	758	4	7	12	17	11		9	4	1	6	1		14	1	1	19		150
A|P				1	1														
A|S				8	2														
C	50		7	12	457	27		11	1	1	12	3		31	2	5	12	1	48
D	10				12	280		40	8	12	2	2		6	4	1	7	18	20
E	48	1			20	42	1	294	3	1	17			3	7	1	24	4	49
F	3								613					1	1		3		1
G		2							2	6	2						2		4
H	40	2		1	29	6		13	1		450	4		22		2	8		265
L						7		1	19	1		221	14	1		27			5
N					1	1		1				10	31		1				2
P	3				16	15	1	2	13	12	1	1		345		2	29		32
Q					8	5		1							40				1
R	3				6							13		1		211	14		3
S	6				48	49		4	26		6			10		1	251		5
T	2				4	2	3								1			45	5
∅	148	1	3	6	136	60		100	32	1	124	9	2	65	12	34	23	6	

Examples

Predicted UCCA STREUSLE	Gold UCCA	
Noun compounds		
tap_water (*unanalyzable*) N.SUBSTANCE	[E tap] [C water]	✗
[P road_construction] N.EVENT	[A road] [P construction]	✗

Examples

Predicted UCCA (STREUSLE)	Gold UCCA	
Noun compounds		
tap_water (*unanalyzable*) N.SUBSTANCE	[E tap] [C water]	✗
[P road_construction] N.EVENT	[A road] [P construction]	✗
Adverbs and linkage		
[H [P Gets_busy]] [L so] [H [P come] [T early]] (Gets_busy: V.VID; come: V.MOTION)	[H [D Gets] [S busy]] [L so] [H [P come] [T early]]	✓✗
[L so] [H [S easy] [A [F to] [P load]]] (load: V.MOTION)	[D [E so] [C easy]] [F to] [P load]	✗

Examples

Predicted UCCA STREUSLE	Gold UCCA	
Noun compounds		
tap_water (*unanalyzable*) N.SUBSTANCE	[E tap] [C water]	✗
[P road_construction] N.EVENT	[A road] [P construction]	✗
Adverbs and linkage		
[H [P Gets_busy]] [L so] [H [P come] [T early]] V.VID, V.MOTION	[H [D Gets] [S busy]] [L so] [H [P come] [T early]]	✓✗
[L so] [H [S easy] [A [F to] [P load]]] V.MOTION	[D [E so] [C easy]] [F to] [P load]	✗
Scene-evoking nouns		
[F a] [C meal] [E [R on] [F the] [C menu]]] N.FOOD, P.LOCUS, N.COMMUNICATION	[F a] [C meal] [E [R on] [F the] [C menu]]]	✓
[P answered] [A [Q all] [A my] [C questions]] V.COMMUNICATION, P.ORIGINATOR, P.GESTALT, N.COMMUNICATION	[P answered] [A [D all] [A my] [P questions]]	✗

Conclusion

- Meaning representations are useful for NLP, and NLG evaluation.

Conclusion

- Meaning representations are useful for NLP, and NLG evaluation.
- They have accurate parsers for various frameworks and languages.

Conclusion

- Meaning representations are useful for NLP, and NLG evaluation.
- They have accurate parsers for various frameworks and languages.
- Linguistic and supervised conversion enable deep comparison.

Thanks!

dh@di.ku.dk

References I

Omri Abend and Ari Rappoport. Universal Conceptual Cognitive Annotation (UCCA). In *Proc. of ACL*, pages 228–238, August 2013. URL http://aclweb.org/anthology/P13-1023.

Wanxiang Che, Longxu Dou, Yang Xu, Yuxuan Wang, Yijia Liu, and Ting Liu. HIT-SCIR at MRP 2019: A unified pipeline for meaning representation parsing via efficient training and effective encoding. In *Proceedings of the Shared Task on Cross-Framework Meaning Representation Parsing at the 2019 Conference on Computational Natural Language Learning*, pages 76 – 85, Hong Kong, China, 2019.

Leshem Choshen and Omri Abend. Inherent biases in reference-based evaluation for grammatical error correction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 632–642, Melbourne, Australia, July 2018a. Association for Computational Linguistics. doi: 10.18653/v1/P18-1059. URL https://www.aclweb.org/anthology/P18-1059.

Leshem Choshen and Omri Abend. Reference-less measure of faithfulness for grammatical error correction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 124–129, New Orleans, Louisiana, June 2018b. Association for Computational Linguistics. doi: 10.18653/v1/N18-2020. URL https://www.aclweb.org/anthology/N18-2020.

Ruixiang Cui and Daniel Hershcovich. Refining implicit argument annotation for UCCA. In *Proc. of DMR*, 2020. URL https://arxiv.org/abs/2005.12889.

Daniel Hershcovich, Omri Abend, and Ari Rappoport. A transition-based directed acyclic graph parser for ucca. In *Proc. of ACL*, pages 1127–1138, 2017. URL http://aclweb.org/anthology/P17-1104.

Daniel Hershcovich, Omri Abend, and Ari Rappoport. Multitask parsing across semantic representations. In *Proc. of ACL*, pages 373–385, 2018. URL http://aclweb.org/anthology/P18-1035.

Daniel Hershcovich, Omri Abend, and Ari Rappoport. Content differences in syntactic and semantic representation. In *Proc. of NAACL-HLT*, pages 478–488, June 2019a. URL https://aclweb.org/anthology/N19-1047.

Daniel Hershcovich, Leshem Choshen, Elior Sulem, Zohar Aizenbud, Ari Rappoport, and Omri Abend. SemEval 2019 task 1: Cross-lingual semantic parsing with UCCA. In *Proc. of SemEval*, 2019b. URL https://aclweb.org/anthology/S19-2001.

References II

Daniel Hershcovich, Miryam de Lhoneux, Artur Kulmizev, Elham Pejhan, and Joakim Nivre. Køpsala: Transition-based graph parsing via efficient training and effective encoding. In *Proceedings of the 16th International Conference on Parsing Technologies and the IWPT 2020 Shared Task on Parsing into Enhanced Universal Dependencies*, pages 236–244, Online, July 2020a. Association for Computational Linguistics. doi: 10.18653/v1/2020.iwpt-1.25. URL https://www.aclweb.org/anthology/2020.iwpt-1.25.

Daniel Hershcovich, Nathan Schneider, Dotan Dvir, Jakob Prange, Miryam de Lhoneux, and Omri Abend. Comparison by conversion: Reverse-engineering UCCA from syntax and lexical semantics. In *Proc. of COLING*, 2020b. URL https://arxiv.org/abs/2011.00834.

Stephan Oepen, Omri Abend, Jan Hajič, Daniel Hershcovich, Marco Kuhlmann, Tim O'Gorman, Nianwen Xue, Jayeol Chun, Milan Straka, and Zdeňka Urešová. MRP 2019: Cross-framework Meaning Representation Parsing. In *Proc. of CoNLL MRP Shared Task*, pages 1 – 27, 2019. URL https://aclweb.org/anthology/K19-2001.pdf.

Stephan Oepen, Omri Abend, Lasha Abzianidze, Johan Bos, Jan Hajič, Daniel Hershcovich, Bin Li, Tim O'Gorman, Nianwen Xue, and Daniel Zeman. MRP 2020: The Second Shared Task on Cross-framework and Cross-Lingual Meaning Representation Parsing. In *Proceedings of the CoNLL 2020 Shared Task: Cross-Framework Meaning Representation Parsing*, pages 1 – 22, Online, 2020.

Hiroaki Ozaki, Gaku Morio, Yuta Koreeda, Terufumi Morishita, and Toshinori Miyoshi. Hitachi at MRP 2020: Text-to-graph-notation transducer. In *Proceedings of the CoNLL 2020 Shared Task: Cross-Framework Meaning Representation Parsing*, pages 40 – 52, Online, 2020.

David Samuel and Milan Straka. ÚFAL at MRP 2020: Permutation-invariant semantic parsing in PERIN. In *Proceedings of the CoNLL 2020 Shared Task: Cross-Framework Meaning Representation Parsing*, pages 53 – 64, Online, 2020.

Elior Sulem, Omri Abend, and Ari Rappoport. Conceptual annotations preserve structure across translations: A French-English case study. In *Proc. of S2MT*, pages 11–22, 2015. URL http://aclweb.org/anthology/W15-3502.

Elior Sulem, Omri Abend, and Ari Rappoport. BLEU is not suitable for the evaluation of text simplification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 738–744, Brussels, Belgium, October-November 2018a. Association for Computational Linguistics. doi: 10.18653/v1/D18-1081. URL https://www.aclweb.org/anthology/D18-1081.

References III

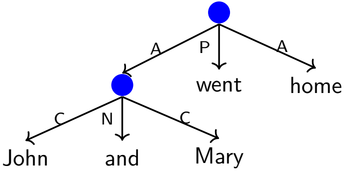
Elior Sulem, Omri Abend, and Ari Rappoport. Semantic structural evaluation for text simplification. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 685–696, New Orleans, Louisiana, June 2018b. Association for Computational Linguistics. doi: 10.18653/v1/N18-1063. URL https://www.aclweb.org/anthology/N18-1063.

Elior Sulem, Omri Abend, and Ari Rappoport. Simple and effective text simplification using semantic and neural methods. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 162–173, Melbourne, Australia, July 2018c. Association for Computational Linguistics. doi: 10.18653/v1/P18-1016. URL https://www.aclweb.org/anthology/P18-1016.

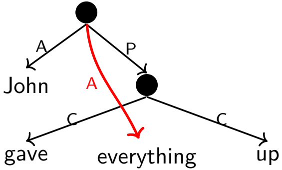
Elior Sulem, Omri Abend, and Ari Rappoport. Semantic structural decomposition for neural machine translation. In *Proc. of *SEM*, 2020.

Structural Properties

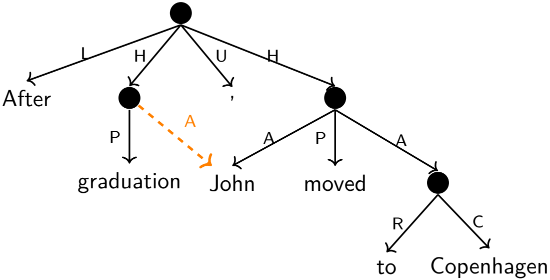
(1) non-terminal nodes



(2) discontinuity



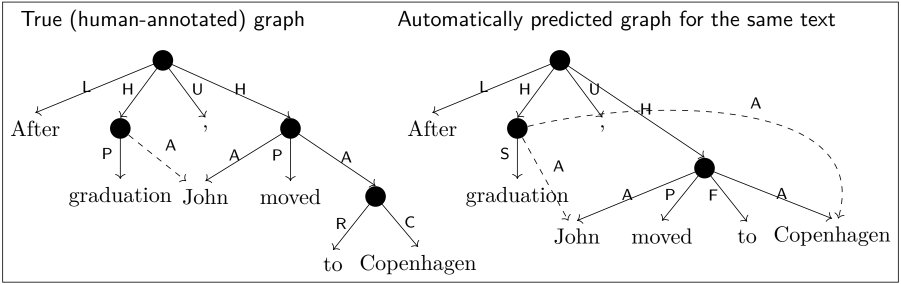
(3) reentrancy



Data Statistics

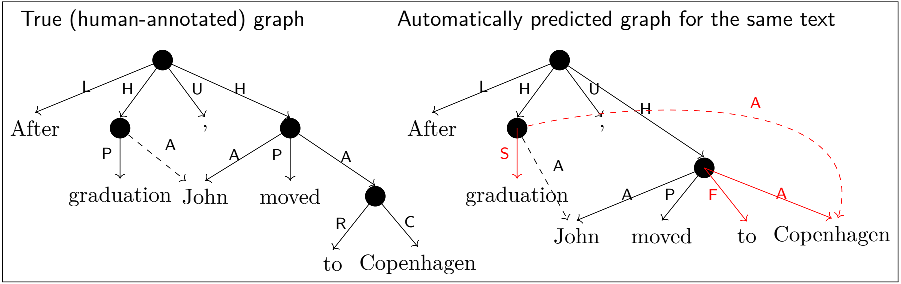
	Wiki	20K			EWT
	en	en	fr	de	en
# sentences	5,141	492	492	6,514	3,520
# tokens	158,739	12,638	13,021	144,529	51,042
# non-terminal nodes	62,002	4,699	5,110	51,934	18,156
% discontinuous	1.71	3.19	4.64	8.87	3.87
% reentrant	1.84	0.89	0.65	0.31	0.83
# edges	208,937	16,803	17,520	187,533	60,739
% primary	97.40	96.79	97.02	97.32	97.32
% remote	2.60	3.21	2.98	2.68	2.68

Evaluation



1. Match primary edges between the graphs by terminal yield and label.
2. Calculate **precision, recall and F1** scores.
3. Repeat for remote edges.

Evaluation



1. Match primary edges between the graphs by terminal yield and label.
2. Calculate **precision, recall and F1** scores.
3. Repeat for remote edges.

Primary

P	R	F1
$\frac{6}{9} = 67\%$	$\frac{6}{10} = 60\%$	64%

Remote

P	R	F1
$\frac{1}{2} = 50\%$	$\frac{1}{1} = 100\%$	67%