SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA

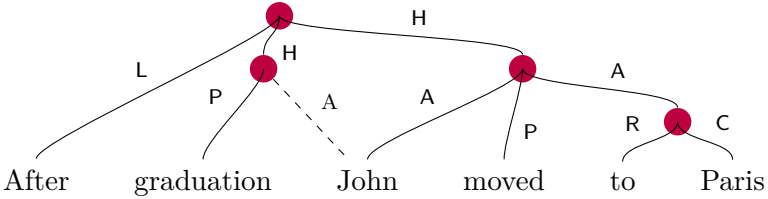
Daniel Hershcovich, Leshem Choshen, Elior Sulem,
Zohar Aizenbud, Ari Rappoport and Omri Abend

האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

June 6, 2019

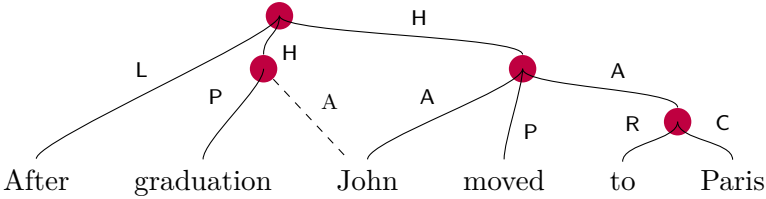
Universal Conceptual Cognitive Annotation (UCCA)

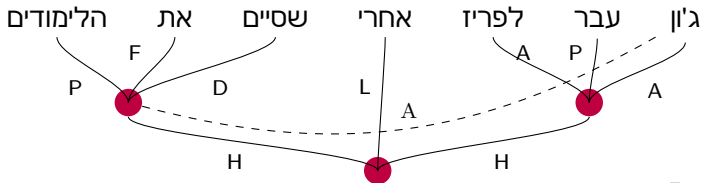
Cross-linguistically applicable semantic representation (Abend and Rappoport, 2013).
Builds on Basic Linguistic Theory (R. M. W. Dixon).
Stable in translation (Sulem et al., 2015).



Universal Conceptual Cognitive Annotation (UCCA)

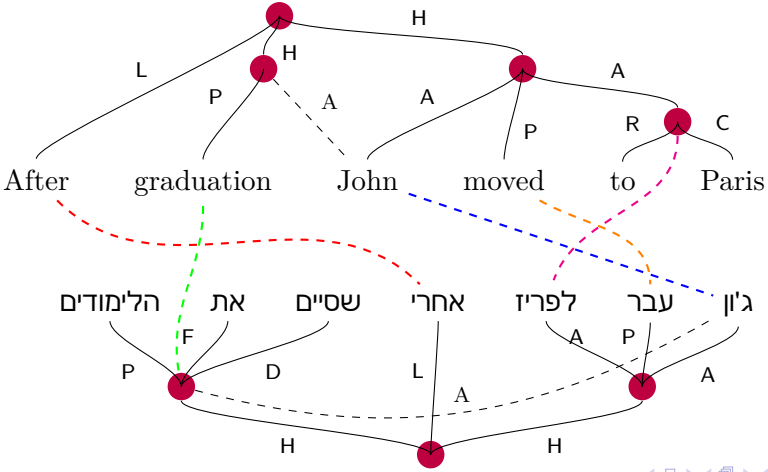
Cross-linguistically applicable semantic representation (Abend and Rappoport, 2013).
Builds on Basic Linguistic Theory (R. M. W. Dixon).
Stable in translation (Sulem et al., 2015).





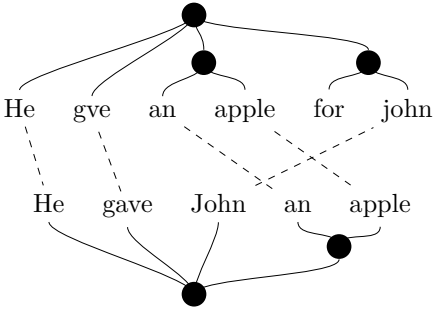
Universal Conceptual Cognitive Annotation (UCCA)

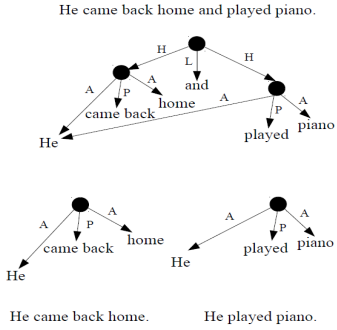
Cross-linguistically applicable semantic representation (Abend and Rappoport, 2013).
Builds on Basic Linguistic Theory (R. M. W. Dixon).
Stable in translation (Sulem et al., 2015).



Applications

- Semantics-based evaluation of
  - Machine translation (Birch et al., 2016)
  - Text simplification (Sulem et al., 2018a)
  - Grammatical error correction (Choshen and Abend, 2018)
- Sentence splitting for text simplification (Sulem et al., 2018b).

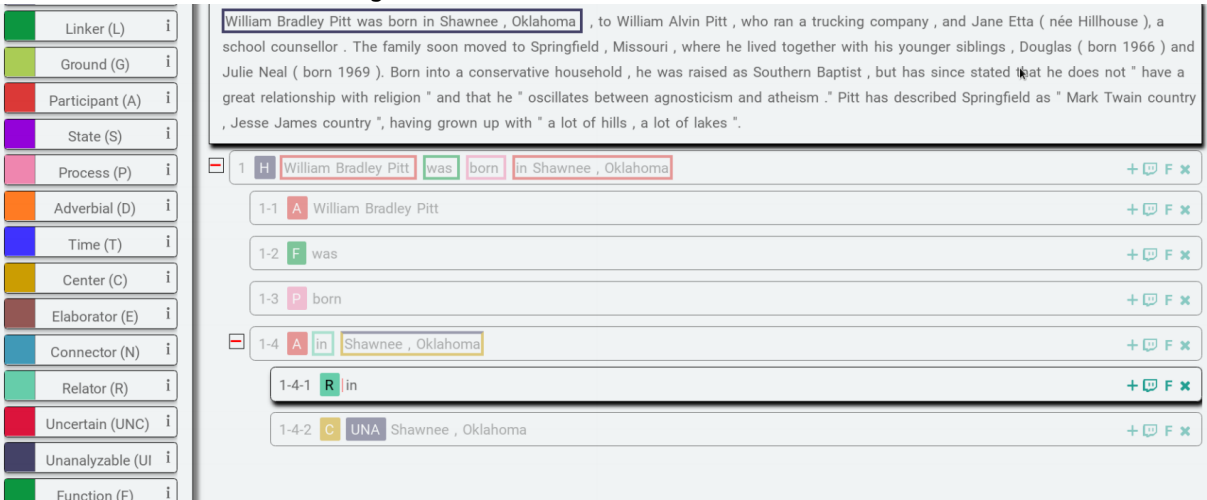




Universal Conceptual Cognitive Annotation (UCCA)

Intuitive annotation interface and guidelines (Abend et al., 2017).

`ucca-demo.cs.huji.ac.il`

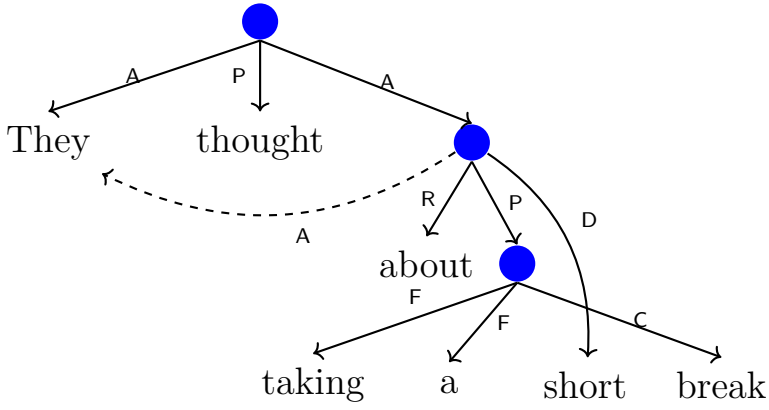


Universal Conceptual Cognitive Annotation (UCCA)

The Task: UCCA parsing in English, German and French in different domains.

Graph Structure

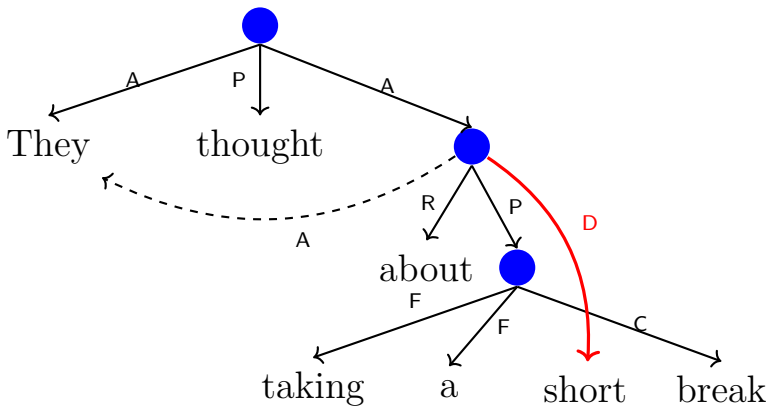
Labeled directed acyclic graphs (DAGs). Complex units are non-terminal nodes.



Graph Structure

Labeled directed acyclic graphs (DAGs). Complex units are non-terminal nodes. Phrases may be discontinuous.

A	Participant
C	Center
D	Adverbial
E	Elaborator
F	Function
G	Ground
H	Parallel scene
L	Linker
P	Process
R	Relator
S	State
U	Punctuation

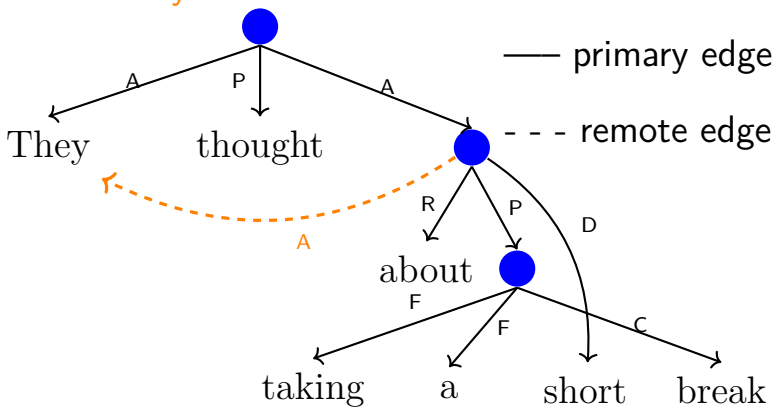


Graph Structure

Labeled directed acyclic graphs (DAGs). Complex units are non-terminal nodes. Phrases may be discontinuous.

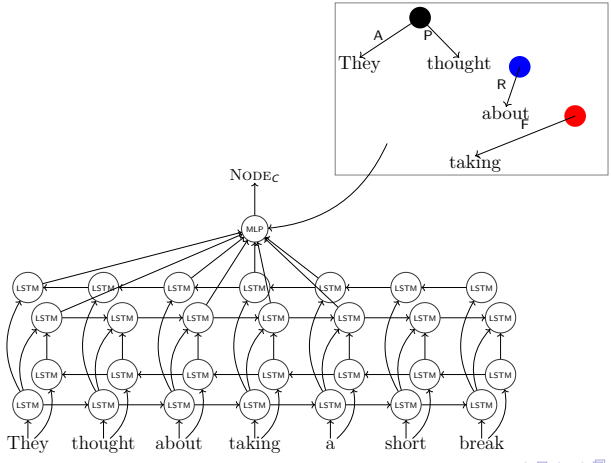
Remote edges enable reentrancy.

A	Participant
C	Center
D	Adverbial
E	Elaborator
F	Function
G	Ground
H	Parallel scene
L	Linker
P	Process
R	Relator
S	State
U	Punctuation



Baseline

TUPA, a transition-based UCCA parser (Hershcovich et al., 2017).
`bit.ly/tupademo`



Data

- English Wikipedia articles (Wiki).
- English-French-German parallel corpus from *Twenty Thousand Leagues Under the Sea* (20K).

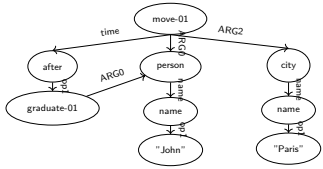
	sentences	tokens
English-Wiki	5,142	158,573
English-20K	492	12,574
French-20K	492	12,954
German-20K	6,514	144,531

Tracks

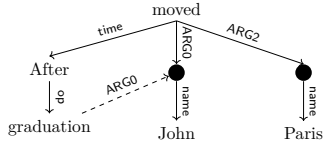
- English {in-domain/out-of-domain} × {open/closed}
- German in-domain {open/closed}
- French *low-resource* (only 15 training sentences)

Conversion

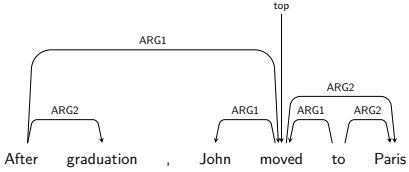
AMR



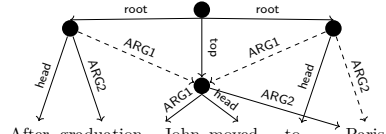
⇔



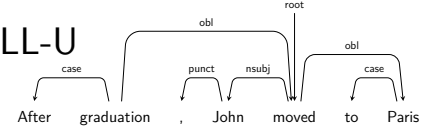
SDP



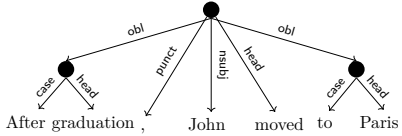
⇔



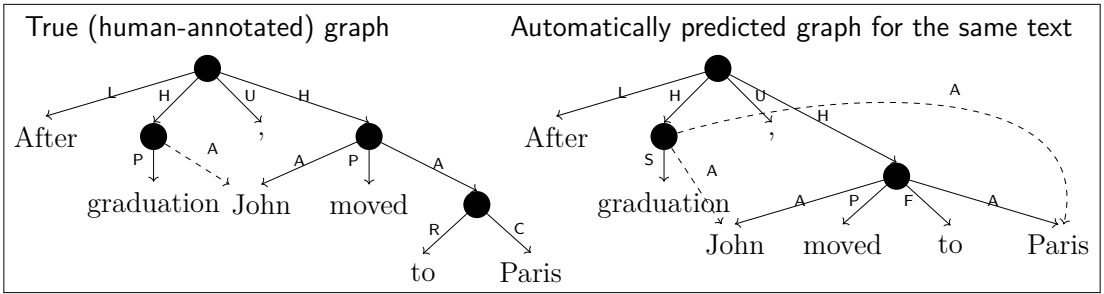
CoNLL-U



⇔

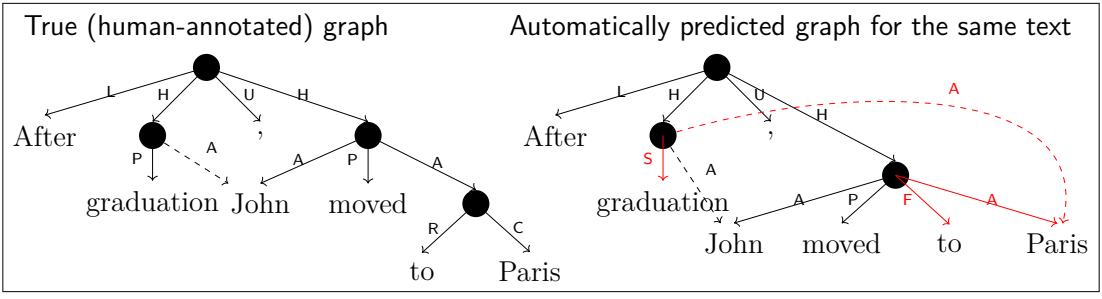


Evaluation



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

Evaluation



1. Match primary edges by terminal yield + 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%

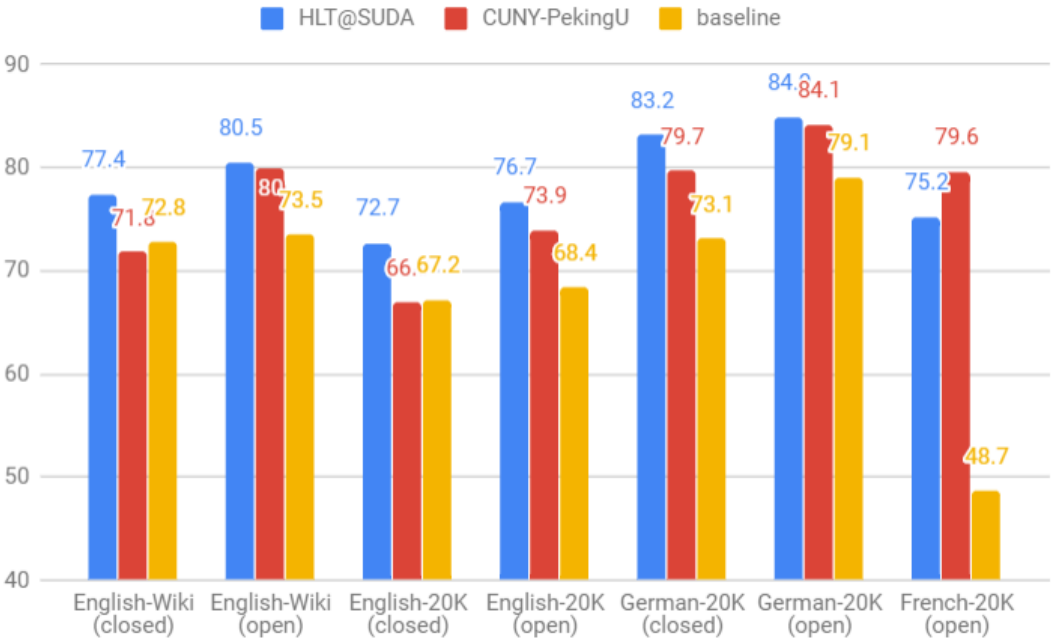
Participating Systems

8 groups in total:

- *MaskParse@Deskiñ* Orange Labs, Aix-Marseille University
- *HLT@SUDA* Soochow University
- *TüPa* University of Tübingen
- *UC Davis* University of California, Davis
- *GCN-Sem* University of Wolverhampton
- *CUNY-PekingU* City University of New York, Peking University
- *DANGNT@UIT.VNU-HCM* University of Information Technology VNU-HCM
- *XLangMo* Zhejiang University

Leaderboard

Track	1st place		2nd place		3rd place		baseline
English-Wiki closed	HLT@SUDA	0.774	baseline	0.728	Davis	0.722	0.728
English-Wiki open	HLT@SUDA	0.805	CUNY-PekingU	0.800	TüPa	0.735	0.735
English-20K closed	HLT@SUDA	0.727	baseline	0.672	CUNY-PekingU	0.669	0.672
English-20K open	HLT@SUDA	0.767	CUNY-PekingU	0.739	TüPa	0.709	0.684
German-20K closed	HLT@SUDA	0.832	CUNY-PekingU	0.797	baseline	0.731	0.731
German-20K open	HLT@SUDA	0.849	CUNY-PekingU	0.841	baseline	0.791	0.791
French-20K open	CUNY-PekingU	0.796	HLT@SUDA	0.752	XLangMo	0.656	0.487


HLT@SUDA
CUNY-PekingU
baseline
90
80
70
60
50
40
77.4
71.8
72.8
80.5
80
73.5
72.7
66
67.2
76.7
73.9
68.4
83.2
79.7
73.1
84.6
84.1
79.1
75.2
79.6
48.7
English-Wiki (closed)
English-Wiki (open)
English-20K (closed)
English-20K (open)
German-20K (closed)
German-20K (open)
French-20K (open)

Main Findings

- HLT@SUDA won 6/7 tracks:
  Neural constituency parser + multi-task + BERT
  French: trained on all languages, with language embedding

Main Findings

- HLT@SUDA won 6/7 tracks:
  Neural constituency parser + multi-task + BERT
  French: trained on all languages, with language embedding
- CUNY-PekingU won the French (open) track:
  TUPA ensemble + synthetic data by machine translation

Main Findings

- HLT@SUDA won 6/7 tracks:
  Neural constituency parser + multi-task + BERT
  French: trained on all languages, with language embedding
- CUNY-PekingU won the French (open) track:
  TUPA ensemble + synthetic data by machine translation

Surprisingly, results in French were close to English and German

- Demonstrates viability of cross-lingual UCCA parsing
- Is this because of UCCA's stability in translation?

Conclusion

- Substantial improvements to UCCA parsing
- High variety of methods
- Successful cross-lingual transfer

Conclusion

- Substantial improvements to UCCA parsing
- High variety of methods
- Successful cross-lingual transfer

Thanks! Annotators, organizers, participants
Daniel Hershcovich, Leshem Choshen, Elior Sulem,
Zohar Aizenbud, Ari Rappoport and Omri Abend

Conclusion

- Substantial improvements to UCCA parsing
- High variety of methods
- Successful cross-lingual transfer

Thanks! Annotators, organizers, participants

Daniel Hershcovich, Leshem Choshen, Elior Sulem,
Zohar Aizenbud, Ari Rappoport and Omri Abend

Please participate in the CoNLL 2019 Shared Task:
Cross-Framework Meaning Representation Parsing
SDP, EDS, AMR and UCCA mrp.nlpl.eu

Evaluation Period: July 8–22, 2019

References I

Abend, O. and Rappoport, A. (2013).
Universal Conceptual Cognitive Annotation (UCCA).
In *Proc. of ACL*, pages 228–238.

Abend, O., Yerushalmi, S., and Rappoport, A. (2017).
UCCAApp: Web-application for syntactic and semantic phrase-based annotation.
Proc. of ACL System Demonstrations, pages 109–114.

Birch, A., Abend, O., Bojar, O., and Haddow, B. (2016).
HUME: Human UCCA-based evaluation of machine translation.
In *Proc. of EMNLP*, pages 1264–1274.

Choshen, L. and Abend, O. (2018).
Reference-less measure of faithfulness for grammatical error correction.
In *Proc. of NAACL (Short papers)*, pages 124–129.

Hershcovich, D., Abend, O., and Rappoport, A. (2017).
A transition-based directed acyclic graph parser for UCCA.
In *Proc. of ACL*, pages 1127–1138.

Sulem, E., Abend, O., and Rappoport, A. (2015).
Conceptual annotations preserve structure across translations: A French-English case study.
In *Proc. of S2MT*, pages 11–22.

Sulem, E., Abend, O., and Rappoport, A. (2018a).
Semantic structural annotation for text simplification.
In *NAACL 2018*, pages 685–696.

Sulem, E., Abend, O., and Rappoport, A. (2018b).
Simple and effective text simplification using semantic and neural methods.
In *Proc. of ACL*, pages 162–173.