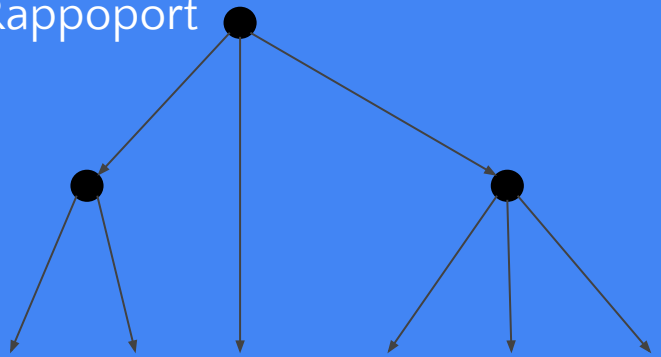
Broad-Coverage Semantic Parsing: A Transition-Based Approach

Daniel Hershcovich, Omri Abend and Ari Rappoport
Hebrew University of Jerusalem



Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing

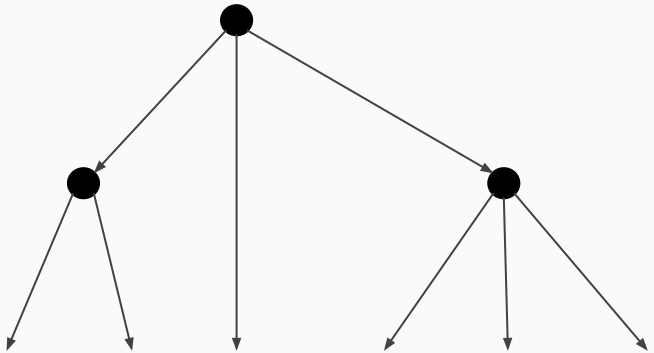
Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing

Given a sequence of tokens $w = w_1, \ldots, w_n$,

A (labeled) directed graph (V, E) where $\{w_i\} \subseteq V$ is a

grounded representation of w.



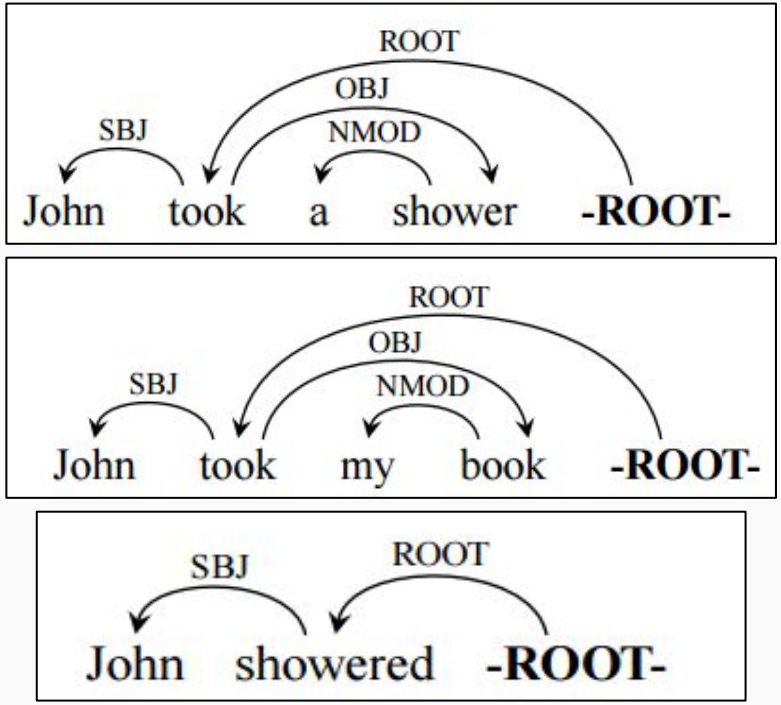
Examples:

Dependency/constituency trees, UCCA…

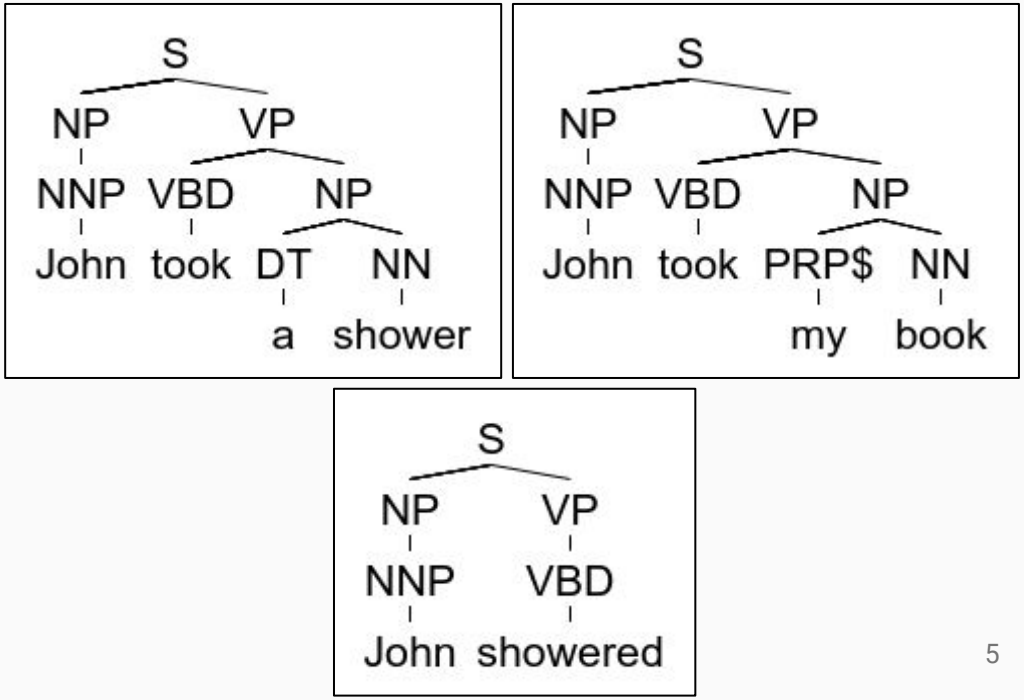
Why Semantic Representation?

Syntactic representation is sensitive to formal variations.

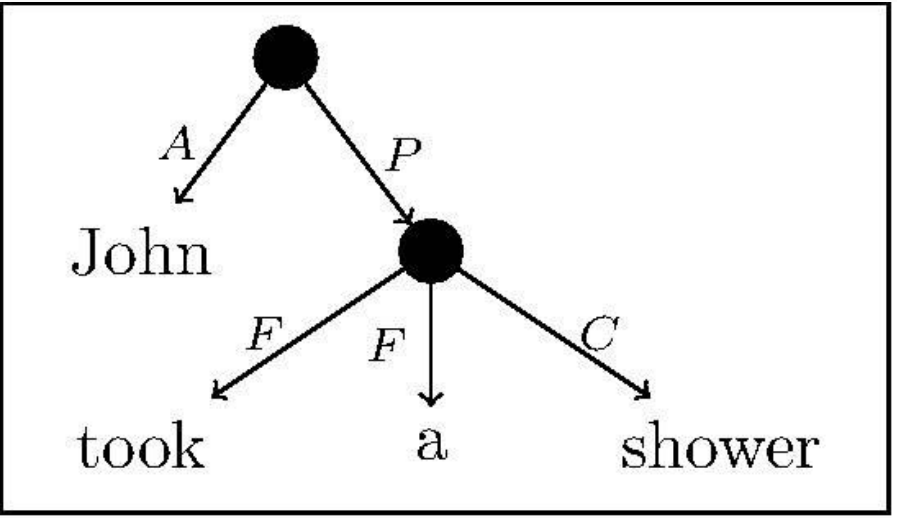
Dependency:

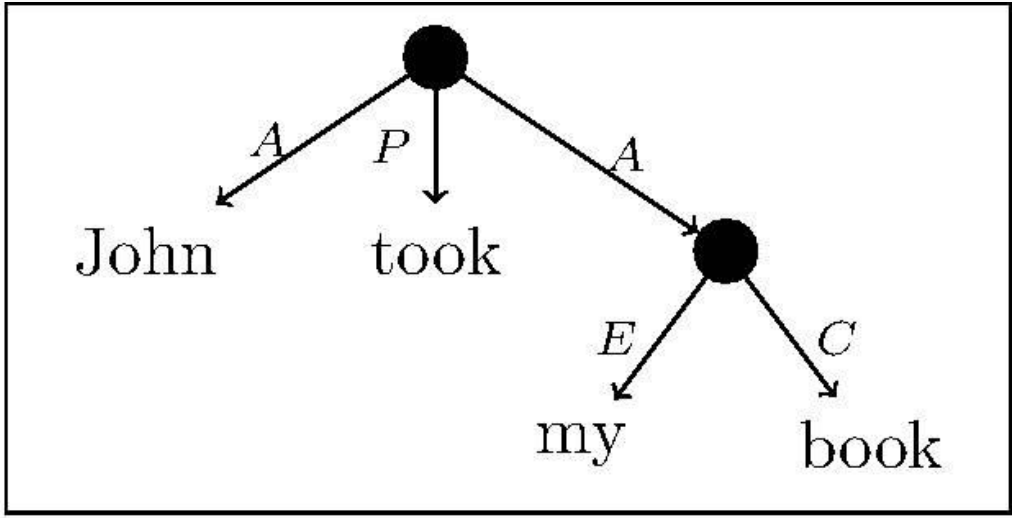


Constituency:

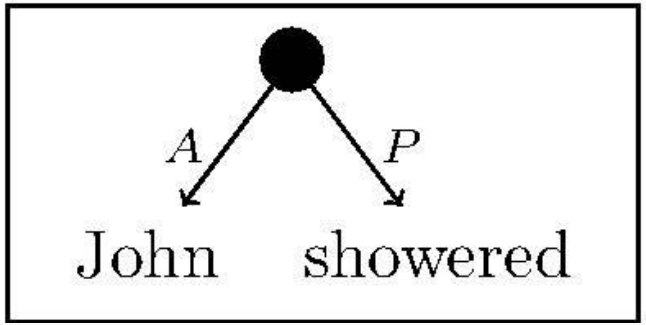


Universal Conceptual Cognitive Annotation (UCCA)





A	participant
P	process
F	function
E	elaborator
C	center

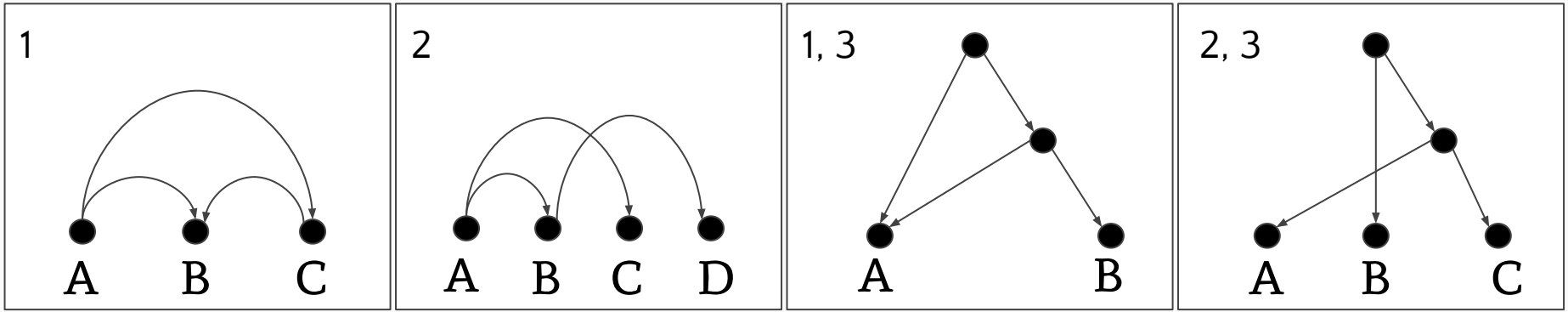


Abend and Rappoport (2013)

Structural Properties

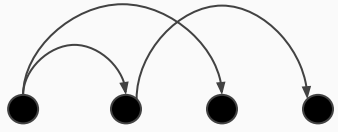
Properties required for full semantic coverage in grounded representations:

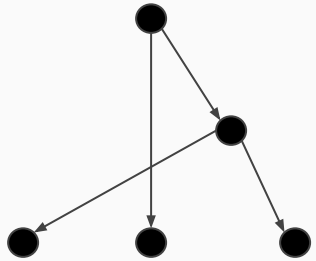
1. **Multiple parents (DAG).**
2. **Non-projectivity (discontinuity).**
3. **Non-terminal nodes.**

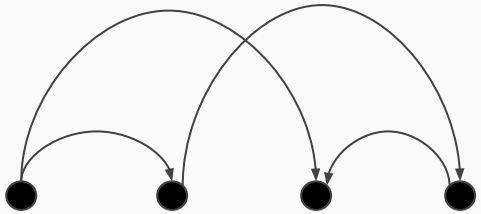


Broad-Coverage Parsing

- Non-projective dependency parsing
- Discontinuous constituency parsing
- Semantic dependency parsing (SDP)





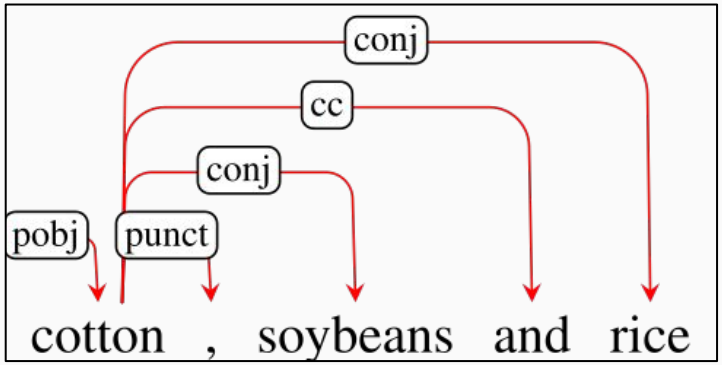


Why Non-Terminals?

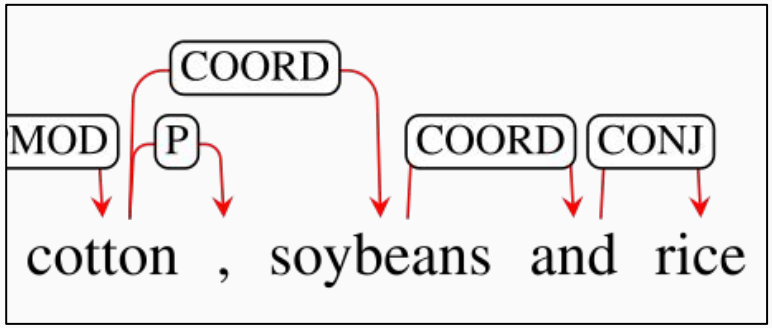
Some frequent constructions do not have one clear head,

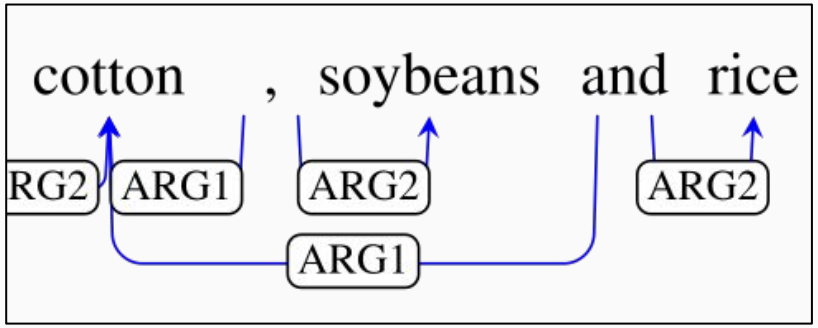
e.g. coordination, some multi-word expressions, compounds.

Coordination in SDP:

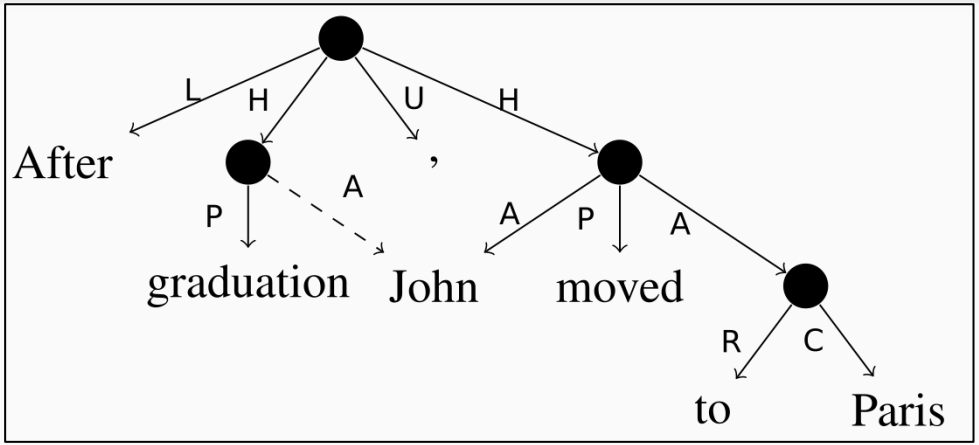


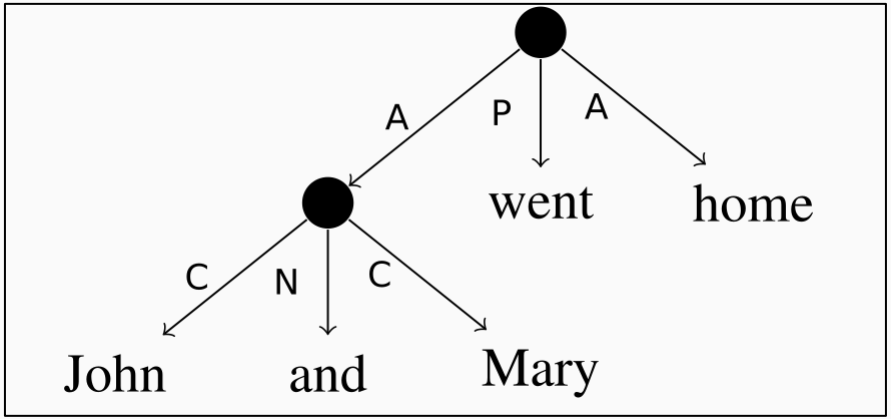
Oepen et al. (2015)



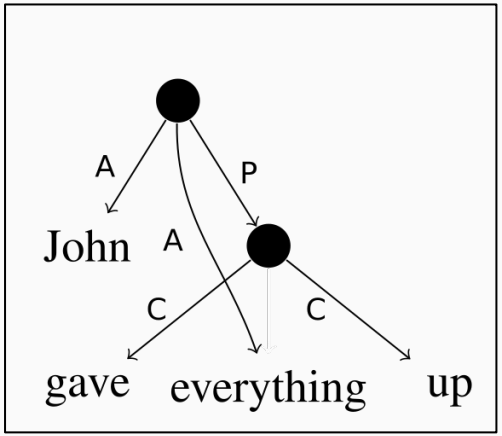


Structural Properties in UCCA

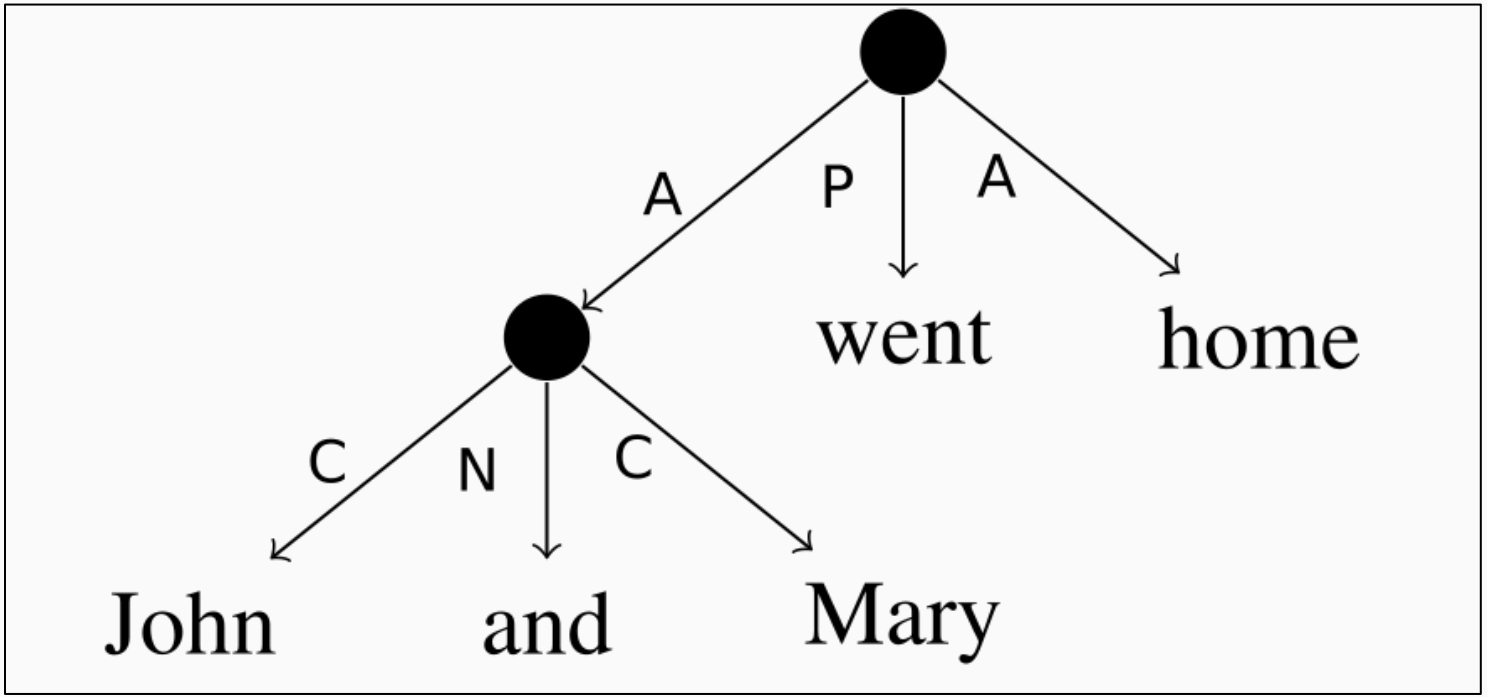




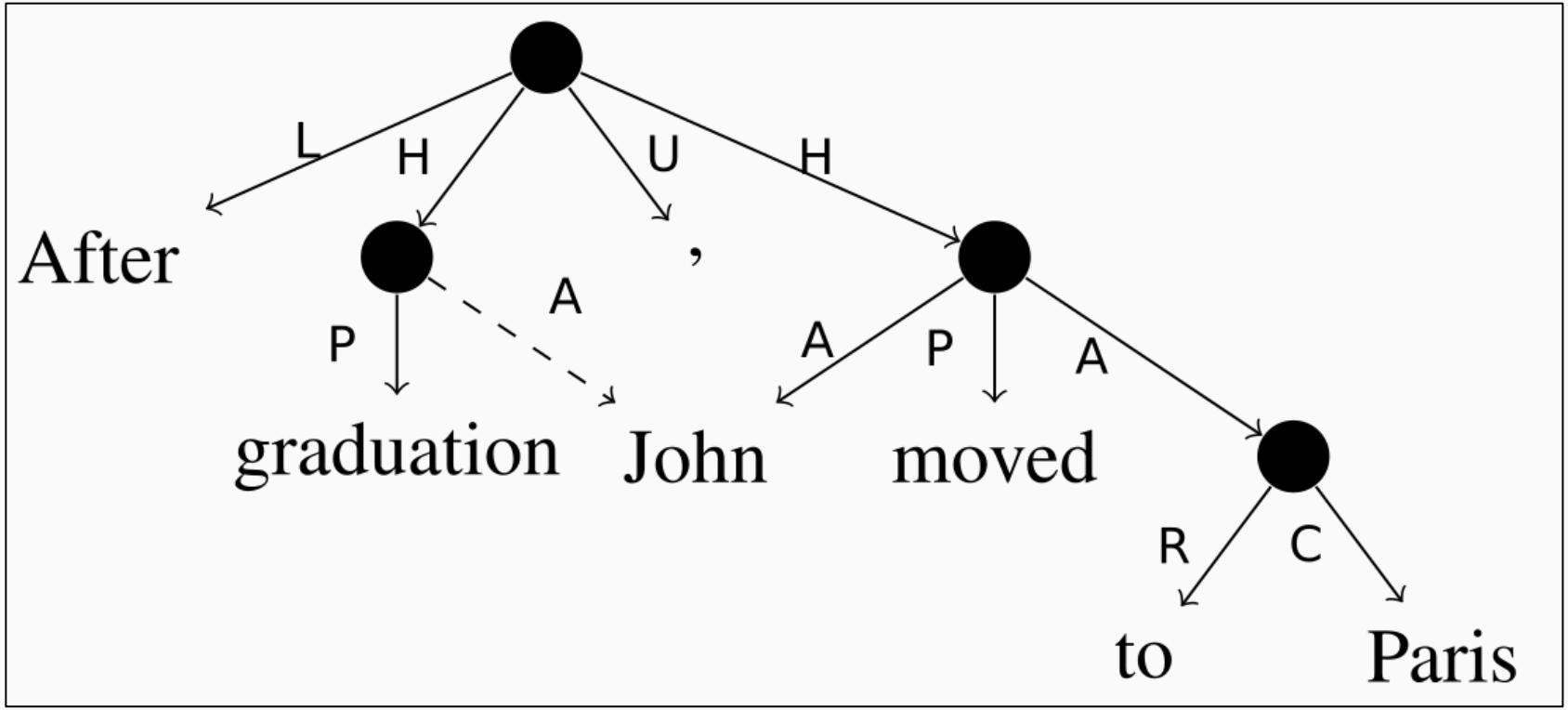
H parallel scene	D adverbial	N connector
A participant	E elaborator	R relator
P process	S state	C center
G ground	L linker	F function



Coordination represented by one parent node



Multiple Parents

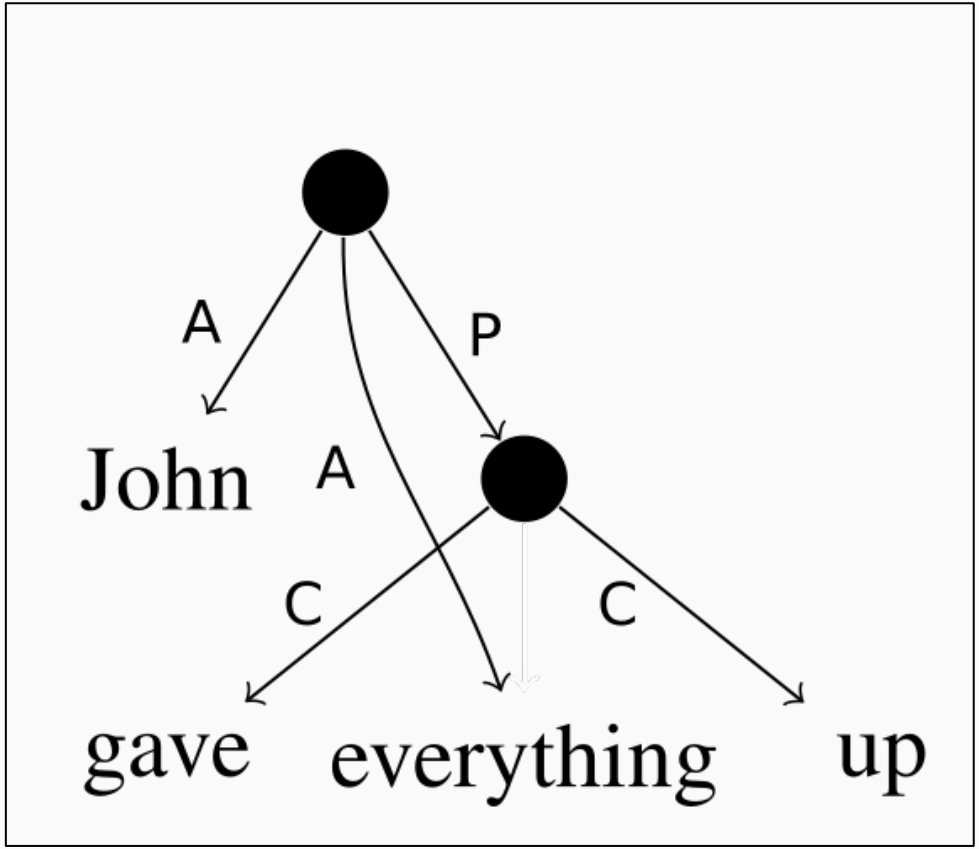


----→ Remote edges denote implicit relations

Discontinuous Units

Multi-word expression

annotated as one unit



Corpus Statistics

160K tokens from English Wikipedia

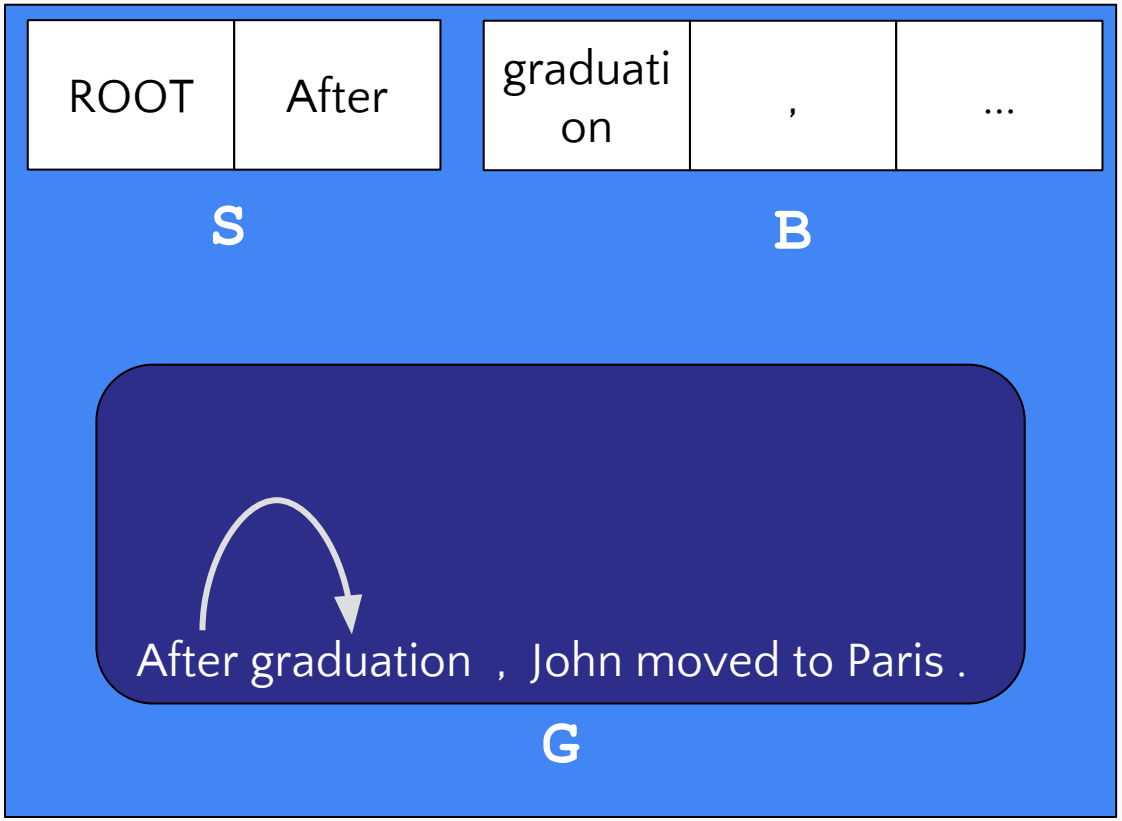
+25K tokens from *Twenty Thousand Leagues Under the Sea*

	Train	Dev	Test	
# passages	281	35	43	154
# sentences	4021	537	608	522
# nodes	277,587	40,700	45,047	29,965
% terminal	42.41	42.8	42.66	41.23
% non-term.	57.59	57.20	57.34	58.77
% discont.	0.52	0.55	0.47	0.79
% >1 parent	2.29	1.89	2.21	1.98

Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing

Transition-Based Dependency Parsing



B:

Buffer of nodes to process, initialized to the list of tokens.

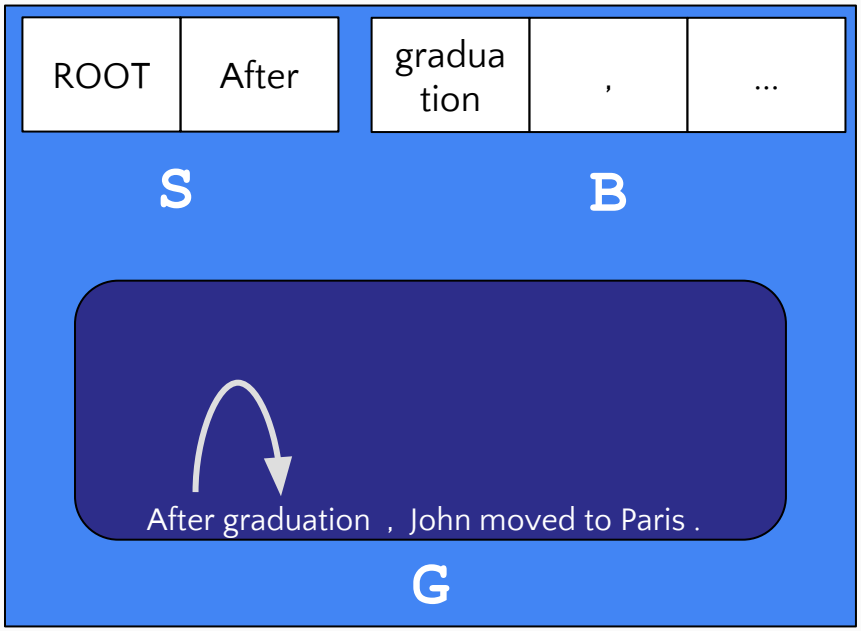
S:

Stack of partially processed nodes, initially just the root.

G:

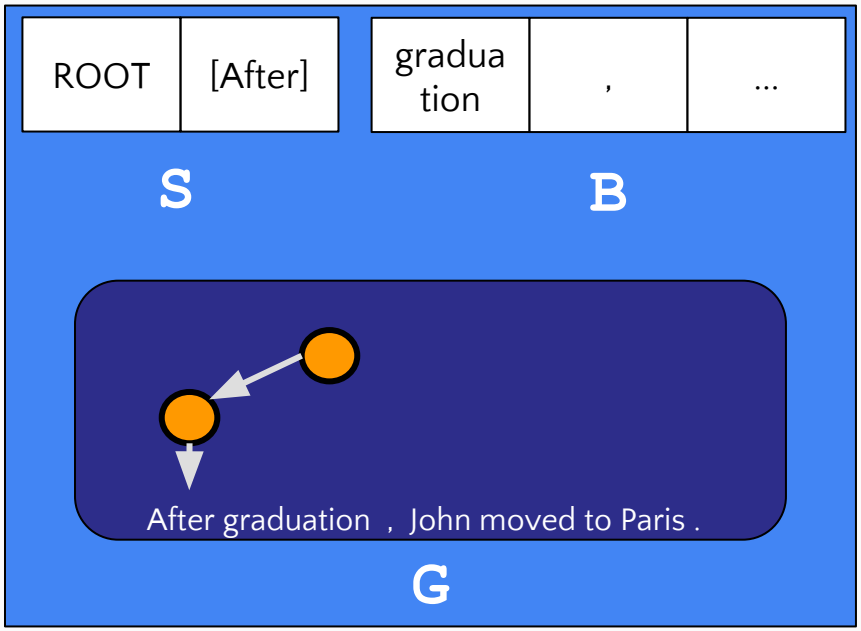
Graph of constructed edges.

Classifier selects next transition given current state



Shift
Reduce
Right-Edge$_X$
Left-Edge$_X$
Swap

Transitions to create new nodes



Shift
Reduce$_X$
Unary$_X$
Finish
Swap

Conversions

1. Convert UCCA to dependency and constituency trees.
2. Apply existing transition-based parsers.
3. Convert back to UCCA.

Dependency parsers:

MaltParser (Nivre 2003), Stack-LSTM Parser (Dyer et al. 2015)

Constituency parser:

UPARSE (Maier 2015): discontinuous constituency parser

Scores on the *Wiki* test set:

	Primary			Remote		
	LP	**LR**	**LF**	**LP**	**LR**	**LF**
Constituency Tree Conversion						
UPARSE	64	67.3	65.4	–	0	0
Upper Bound	100	100	100	–	0	0
Dependency Tree Conversion						
$\text{Malt}_{\text{arc-standard}}$	63.4	57.3	60.1	–	0	0
$\text{Malt}_{\text{arc-eager}}$	63.9	57.9	60.5	–	0	0
LSTM	**73.2**	**66.2**	**69.2**	–	0	0
Upper Bound	93.8	83.7	88.4	–	0	0

Upper bound is due to lossy conversion algorithms.

Classifiers

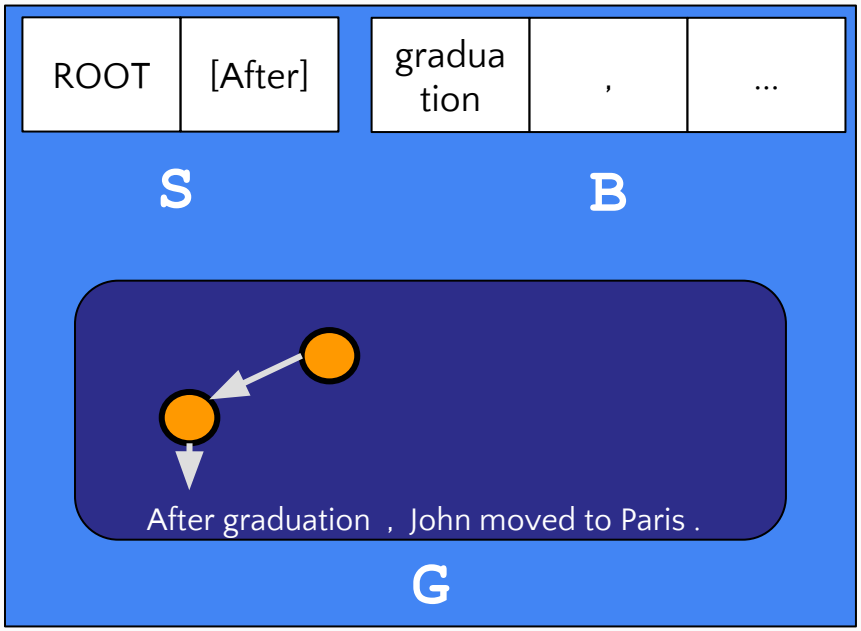
- MaltParser: perceptron/SVM
- UPARSE: perceptron
- Stack-LSTM parser: recurrent neural network

  \+ continuous features

Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing

Discontinuous DAG parser



Perceptron

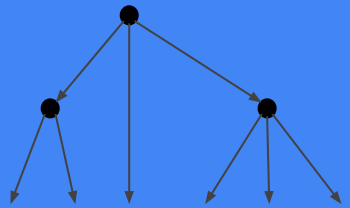
Shift
Reduce
Node$_X$
Left-Edge$_X$
Right-Edge$_X$
Left-Remote$_X$
Right-Remote$_X$
Swap
Finish

Scores on the *Wiki* test set and the *20K leagues* set:

	Primary			Remote		
	LP	**LR**	**LF**	**LP**	**LR**	**LF**
BSP	62.4	56	59	15.3	11.8	13.3
BSP_{Tree}	63.8	56.5	59.9	–	0	0
Out-of-domain						
BSP	60.6	53.9	57.1	20.2	10.3	13.6
BSP_{Tree}	60.2	52.8	56.2	–	0	0

(BSP_{Tree} trained on converted trees without remote edges)

Conclusion



- The structural desiderata of grounded semantic parsing is not supported by today's parsers
- We present a transition-based system that does
- Encouraging results with UCCA suggest that NN-based classification may be helpful for better performance

Future Work

- Neural network for BSP classifier
- Improved conversions
- Beam search
- More languages, e.g. German

Thank you

References

- Angelina Ivanova et al. Who did what to whom?: A contrastive study of syntacto-semantic dependencies. 2012. In *Proc. of LAW*.
- Stephan Oepen et al. SemEval 2015 Task 18: Broad-Coverage Semantic Dependency Parsing. 2015. In *Proc. of SemEval*.
- Omri Abend and Ari Rappoport. Universal Conceptual Cognitive Annotation (UCCA). In *Proc. of ACL*.