Deep Learning in NLP: Neural Semantic Role Labeling

Christian Wurm & Tatiana Bladier

Heinrich-Heine-Universität Düsseldorf
Wintersemenster 2018/2019

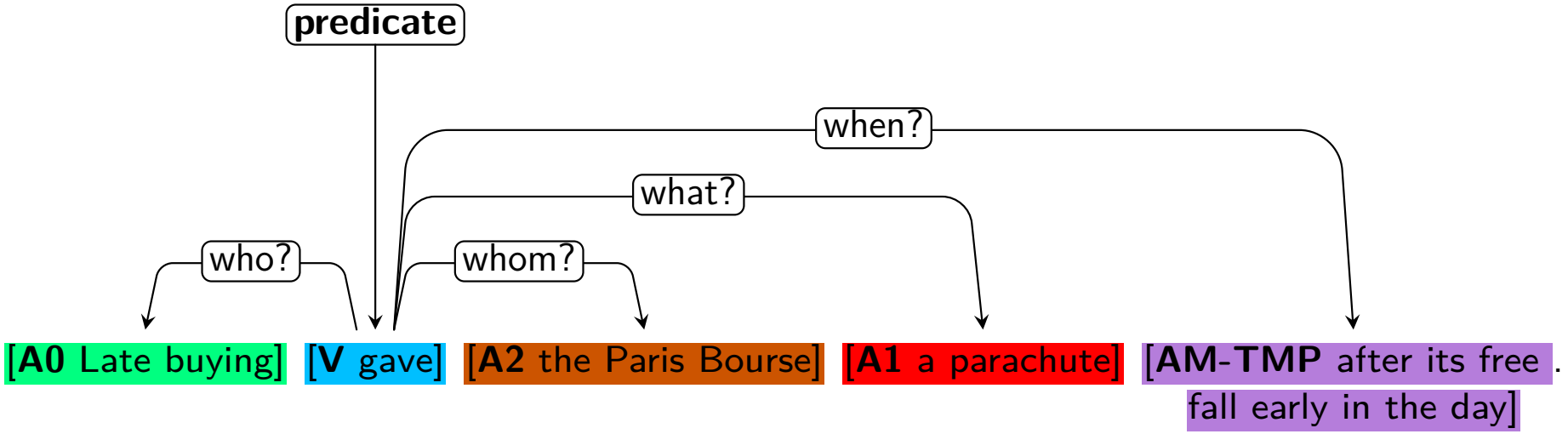
Overview

1. Semantic Role Labeling (SRL)
2. Neural SRL: Syntax-agnostic
3. Neural SRL: Syntax-aware

Semantic Role Labeling: WHO did WHAT to WHOM?

Semantic Role Labeling is a task of:

- identifying semantic arguments of a predicate (WHO did WHAT to WHOM, WHERE, WHEN, WHY etc.)
- and labeling the arguments with their semantic roles (e.g. *Experiencer*, *Content* or A_0, A_1 etc.)



Semantic Role Labeling: From sentences to propositions

Compare different wordings:

Mr. White met Mr. Williams.

Mr. White had a meeting with Mr. Williams.

Mr. White and Mr. Williams met.

Mr. White and Mr. Williams were in a meeting.

$\Downarrow$

Proposition: meet (Mr. White , Mr. Williams)

$\Downarrow$

Information Extraction, Question Answering, Text Summarization, Reasoning

Proposition Bank (PropBank)

PropBank:

- ⋆ Corpus of text annotated with information about basic semantic propositions (Kingsbury and Palmer, 2002; Palmer et al., 2005)
- ⋆ `https://propbank.github.io/`
- ⋆ Searchable frame files: `http://verbs.colorado.edu/propbank/framesets-english-aliases/`

PropBank frame examples

PropBank frames for *give*:

`http://verbs.colorado.edu/propbank/framesets-english-aliases`

give.01 "transfer"	
Roles:	Arg0-PAG: giver (agent) Arg1-PPT: thing given (theme) Arg2-GOL: entity given to (recipient)
Example:	*The executives gave the chefs a standing ovation.*

give.14 "to be likely to, given to"	
Roles:	Arg1-PAG: Person/entity likely to do/be a certain way Arg2-PRD: Thing which Arg1 is likely to do/be
Example:	*One is given to wonder how the police could have missed the very blunt instrument they claimed they were looking for.*

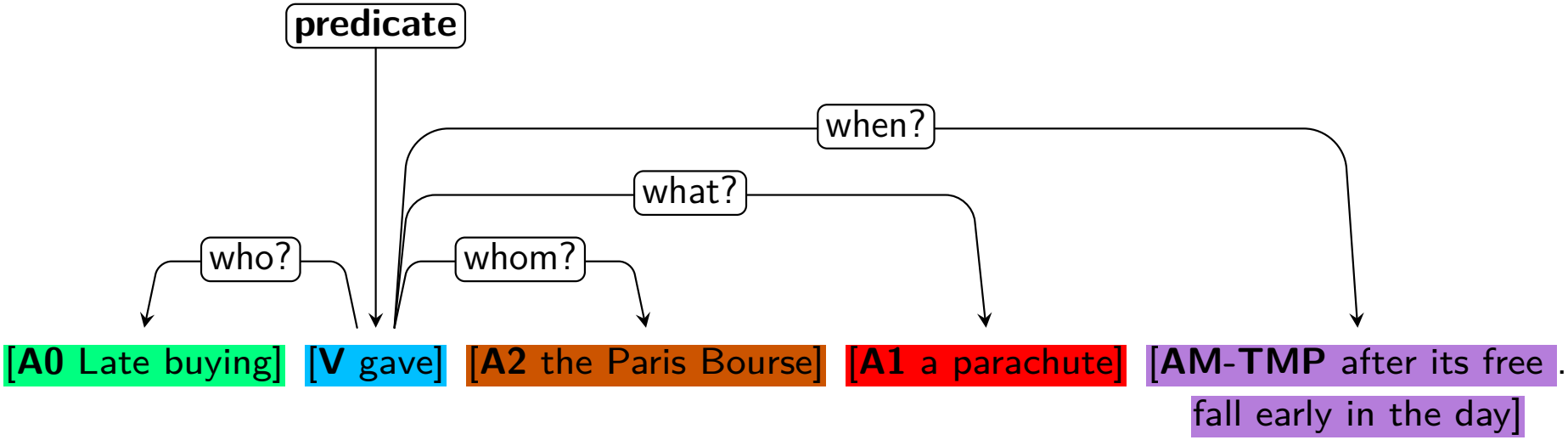
PropBank Semantic Role Labels

Why numbered arguments (Palmer et al., 2013):

- ⋆ Lack of consensus concerning semantic role labels.
- ⋆ Numbers correspond to verb-specific labels.
- ⋆ Arg0 – Proto-Agent, Arg1 – Proto-Patient.
- ⋆ Arg2-5 are highly variable and overloaded – poor performance.

PropBank: Core Semantic Roles (Palmer et al., 2013)

Arg0	agent, experiencer
Arg1	patient, theme
Arg2	benefactive / instrument / attribute / end state
Arg3	start point / benefactive / instrument /attribute
Arg4	end point



PropBank: Argument Modifiers (Palmer et al., 2013)

AM-TMP	when? *on Sunday, tomorrow, recently*
AM-LOC	where? *in New York, at home, in the kitchen*
AM-DIR	where to / from? *from Germany, to the lighthouse*
AM-MNR	how? *seriously, quickly, intentionally*
AM-PRP/CAU	why? *due to, because of*
AM-REC	reciprocal *herself, one another*
AM-GOL	end point of motion verbs *to the floor, to Mary*
AM-PRD	secondary predication *eat an apple unwashed*
AM-ADV	miscellaneous (nothing else fits)

How do the files in PropBank look like?

WORDS---->	NE--->	POS	PARTIAL_SYNT		FULL_SYNT------>	VS	TARGETS	PROPS------->	
The	*	DT	(NP*	(S*	(S(NP*	-	-	(A0*	(A0*
$	*	$	*	*	(ADJP(QP*	-	-	*	*
1.4	*	CD	*	*	*	-	-	*	*
billion	*	CD	*	*	*))	-	-	*	*
robot	*	NN	*	*	*	-	-	*	*
spacecraft	*	NN	*)	*	*)	-	-	*)	*)
faces	*	VBZ	(VP*)	*	(VP*	01	face	(V*)	*
a	*	DT	(NP*	*	(NP*	-	-	(A1*	*
six-year	*	JJ	*	*	*	-	-	*	*
journey	*	NN	*)	*	*	-	-	*	*
to	*	TO	(VP*	(S*	(S(VP*	-	-	*	*
explore	*	VB	*)	*	(VP*	01	explore	*	(V*)
Jupiter	(ORG*)	NNP	(NP*)	*	(NP(NP*)	-	-	*	(A1*
and	*	CC	*	*	*	-	-	*	*
its	*	PRP$	(NP*	*	(NP*	-	-	*	*
16	*	CD	*	*	*	-	-	*	*
known	*	JJ	*	*	*	-	-	*	*
moons	*	NNS	*)	*)	*))))))))	-	-	*)	*)
.	*	.	*	*)	*)	-	-	*	*

Other SRL Resources

Two other major frequently used SRL resources:

- ⋆ FrameNet
  `https://framenet.icsi.berkeley.edu/fndrupal/`
- ⋆ VerbNet
  `https://verbs.colorado.edu/~mpalmer/projects/verbnet.html`

Non-verbal predicates

for.01 "in favor of, supporting"	
Roles:	Arg1-PAG: advocate, supporter Arg2-PPT: in favor of what
Example:	*I'm all for a bit of protectionism when it comes to labour markets.*

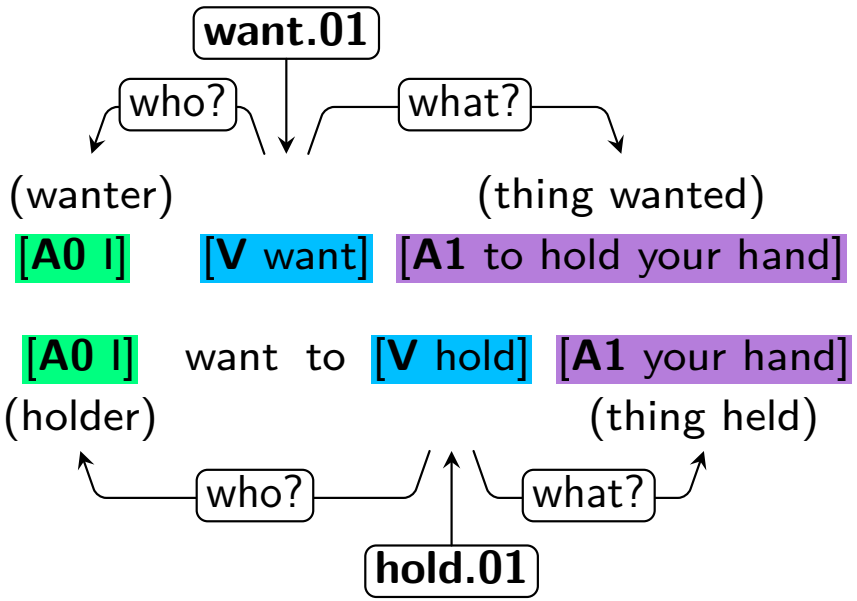
friendly.01 "friendly"	
Roles:	Arg1-PPT: friendly entity Arg2-GOL: target of friendship
Example:	She is friendly with my cousins.

http://verbs.colorado.edu/propbank/framesets-english-aliases

General SRL pipeline (1)

Two steps pipeline (Marcheggiani et al., 2013):

- ⋆ Identify words that are predicates.
- ⋆ Identify and label all the arguments for each predicate.
- ⋆ If a sequence has n_p predicates, the sequence is processed n_p times.



General SRL pipeline (2) (Marcheggiani et al., 2013)

Argument identification:

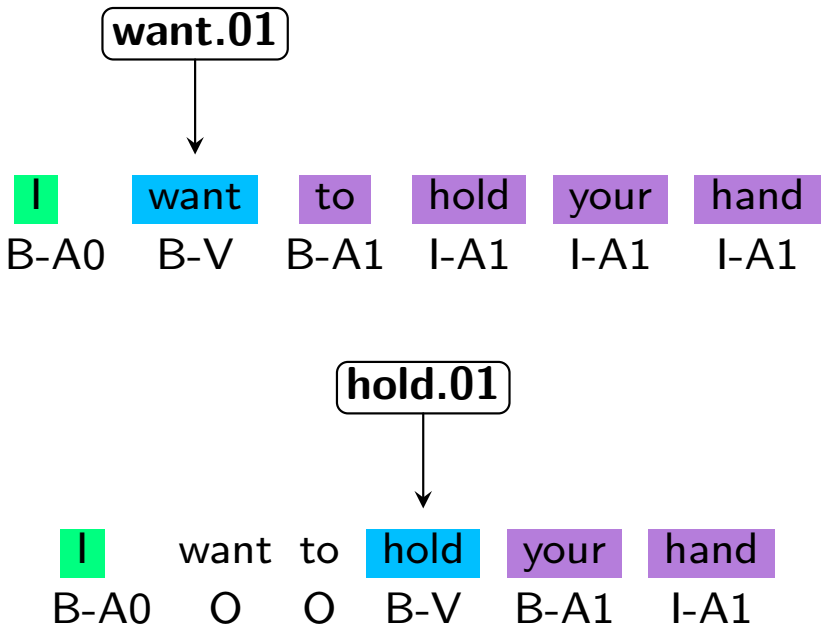
- ⋆ Hand-crafted rules on the full syntactic trees (Xue and Palmer, 2004).
- ⋆ Binary classifier (Pradhan et al., 2005).
- ⋆ Both (Punyakanok et al., 2008).

Role labeling + Inference:

- ⋆ Labeling is performed using a classifier (SVM, logistic regression).
- ⋆ Most of the features are syntactic (Gildea and Jurafsky, 2002).
- ⋆ For each argument we get a label distribution.
- ⋆ Argmax over roles will result in a local assignment.
- ⋆ No guarantee the labeling is well formed
  $\rightarrow$ overlapping arguments, duplicate core roles etc.
- ⋆ Toutanova et al. (2008); Täckström et al. (2015); Björkelund et al. (2010)

Neural syntax-agnostic SRL methods (1)

- ⋆ SRL as a sequence labeling problem.
- ⋆ Argument identification and role labeling in one step.
- ⋆ BIO-scheme for the argument labeling.



(Marcheggiani et al., 2013)

Neural syntax-agnostic SRL methods (2)

General architecture:

- ⋆ Word encoding.
- ⋆ Sentence encoding (e.g. via LSTM).
- ⋆ Decoding.
- ⋆ No use of any kind of treebank syntax (not trivial to encode).

(Marcheggiani et al., 2013)

End-to-end learning of SRL using RNNs

Model proposed by Zhou and Xu (2015):

http://www.aclweb.org/anthology/P15-1109

https://github.com/sanjaymeena/semantic_role_labeling_deep_learning

- ⋆ No syntax information.
- ⋆ Minimal word representation.
- ⋆ Sentence encoding with BiLSTM.
- ⋆ Conditional Random Fields (CRF) layer for label prediction:

  → CRF focus on sentence level instead of individual positions for sequence predictions;
  → CRFs can produce higher tagging accuracy in general (Huang et al., 2015).

BiLSTM-CRF Model (Huang et al., 2015)

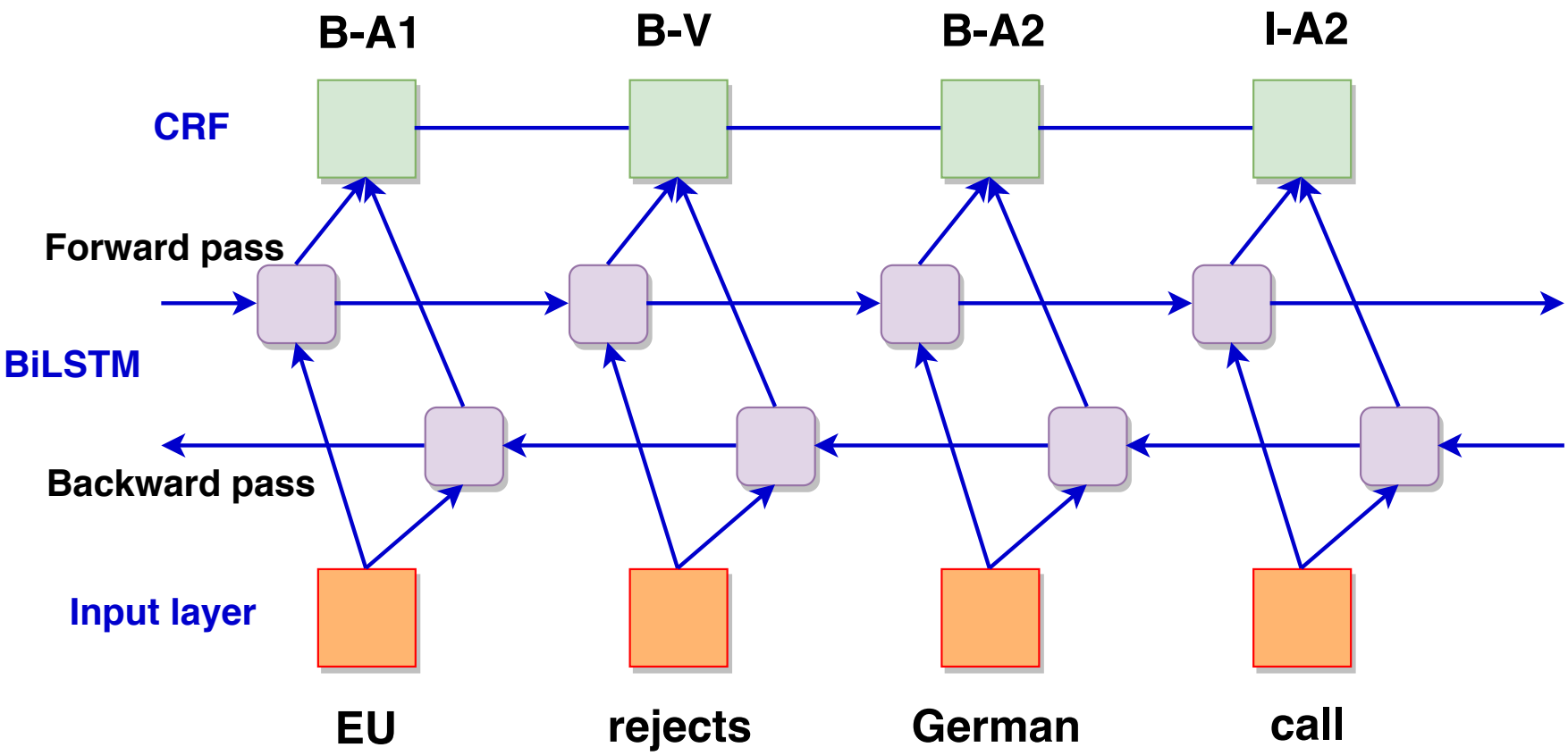


Figure: BiLSTM-CRF Model (Huang et al., 2015)

BiLSTM-CRF Model for SRL (Zhou and Xu, 2015)

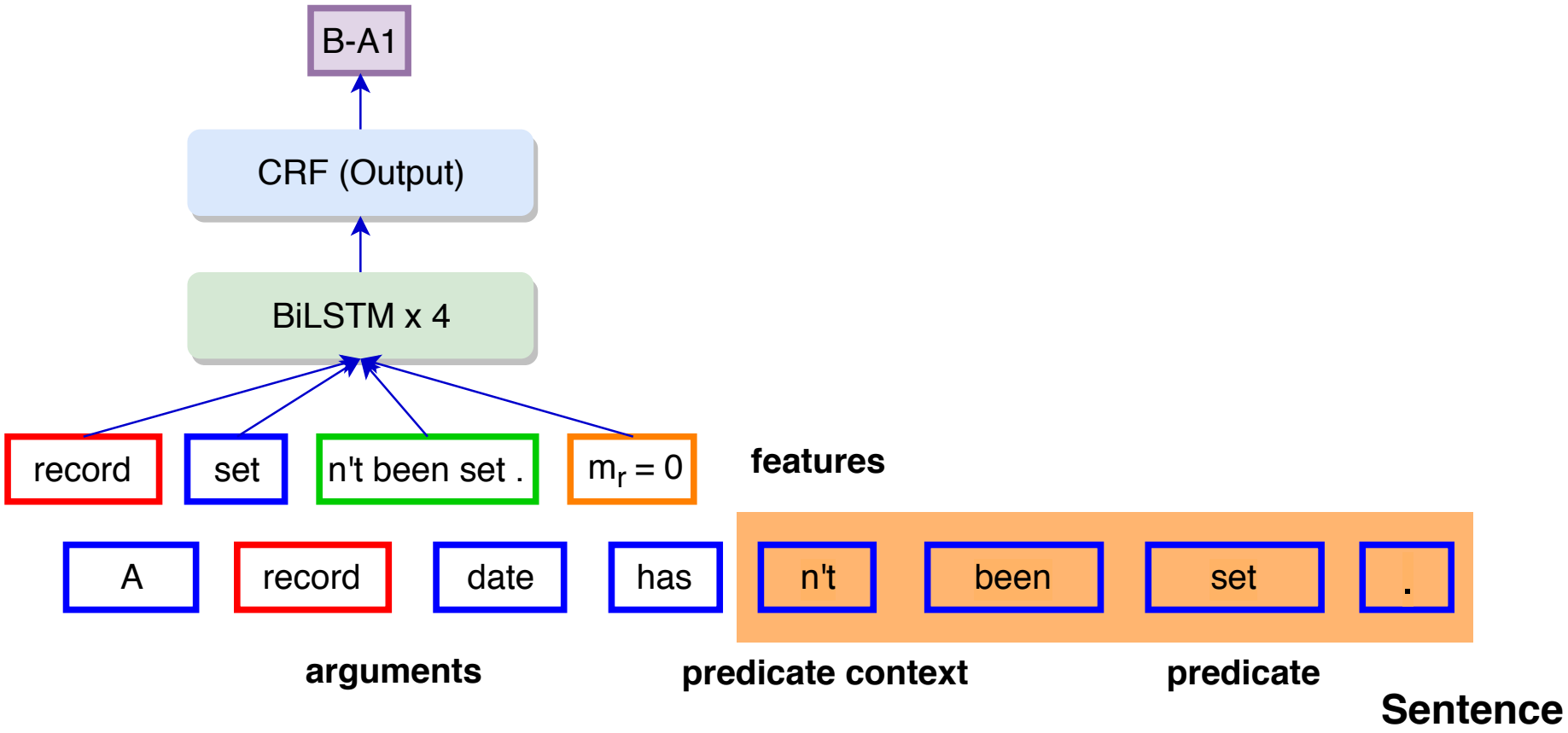


Figure: BiLSTM-CRF for Semantic Role Labeling (Zhou and Xu, 2015)

BiLSTM-CRF SRL Model: Features

token	arg	pred	context-pred	m_r	label
1	A	set	n't been set .	0	B-A1
2	record	set	n't been set .	0	I-A1
3	date	set	n't been set .	0	I-A1
4	has	set	n't been set .	0	O
5	n't	set	n't been set .	1	B-AM-NEG
6	been	set	n't been set .	1	O
7	set	set	n't been set .	1	B-V
8	.	set	n't been set .	1	O

Table: Features used in Zhou and Xu (2015)

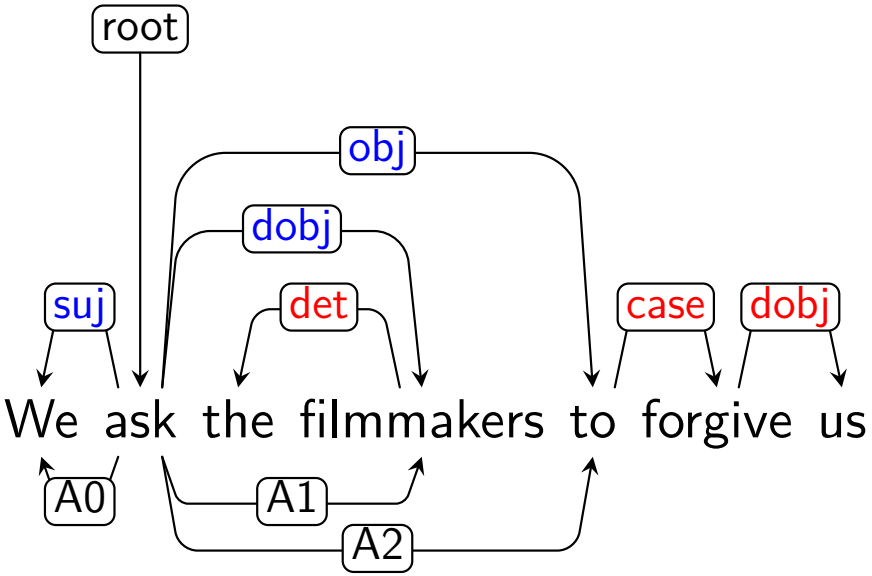
Neural syntax-aware SRL methods (Marcheggiani et al., 2013)

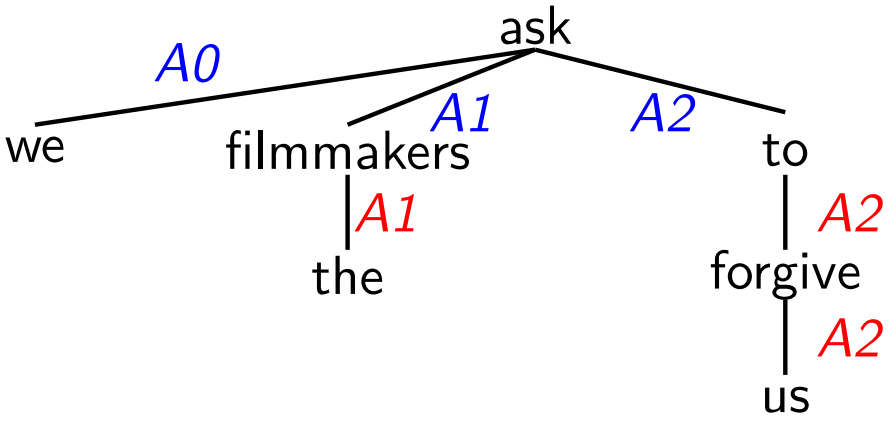
- ⋆ Propbank-style SRL formalism is closely tied to syntax: 98.7% of SRL arguments match an unlabeled constituent in the gold syntax tree (He et al., 2017).
- ⋆ Do Deep SR Labeler inherently learn syntax?

- ⋆ Syntax is not trivial to encode → different approaches.
- ⋆ POS tags are beneficial for SRL (Marcheggiani et al., 2017).
- ⋆ Gold syntax is beneficial, but hard to encode (He et al., 2017).
- ⋆ Performance of the SRL labeler depends on the performance of the underlying syntactic parser (He et al., 2017).

SRL + Graph Convolutional Networks

- ⋆ Model proposed by Marcheggiani and Titov (2017).
- ⋆ Based on encoding of syntactic dependencies with Graph Convolutional Networks (GCNs).
- ⋆ GCNs are multilayer neural networks operating on graphs (Kipf and Welling, 2016) → applicable to dependency graphs.





Syntactic GCN Example (Marcheggiani and Titov, 2017)

- ⋆ Graph Convolutional Networks encode for every node in the graph relevant information about its neighborhood as a real-valued feature vector.

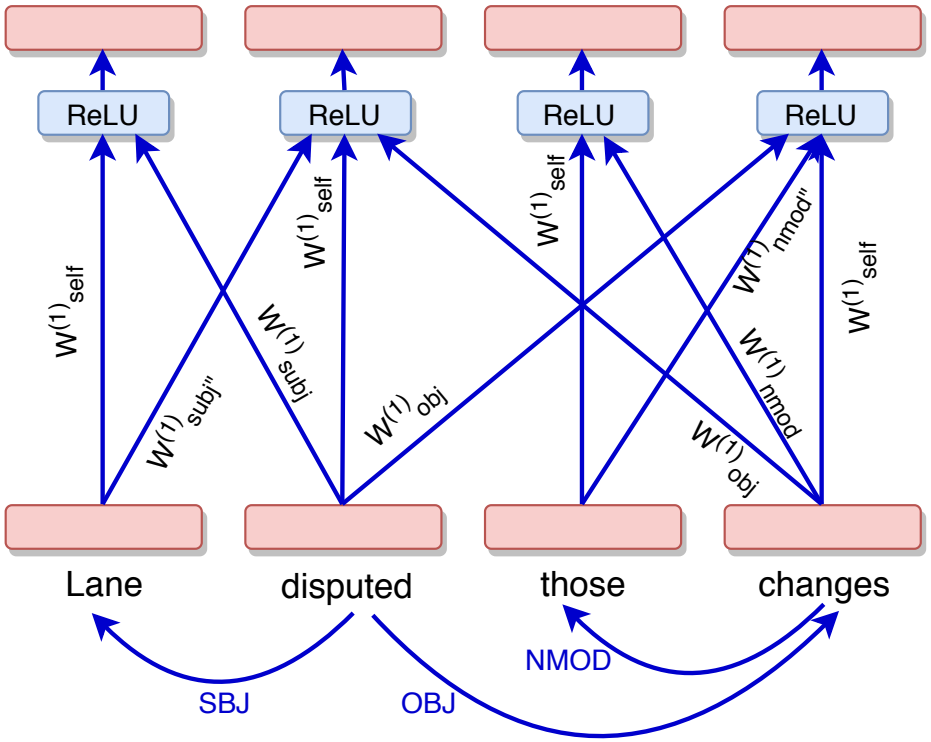


Figure: Syntactic GCN Example (Marcheggiani and Titov, 2017)

LSTM-GCN Model for SRL (Marcheggiani and Titov, 2017)

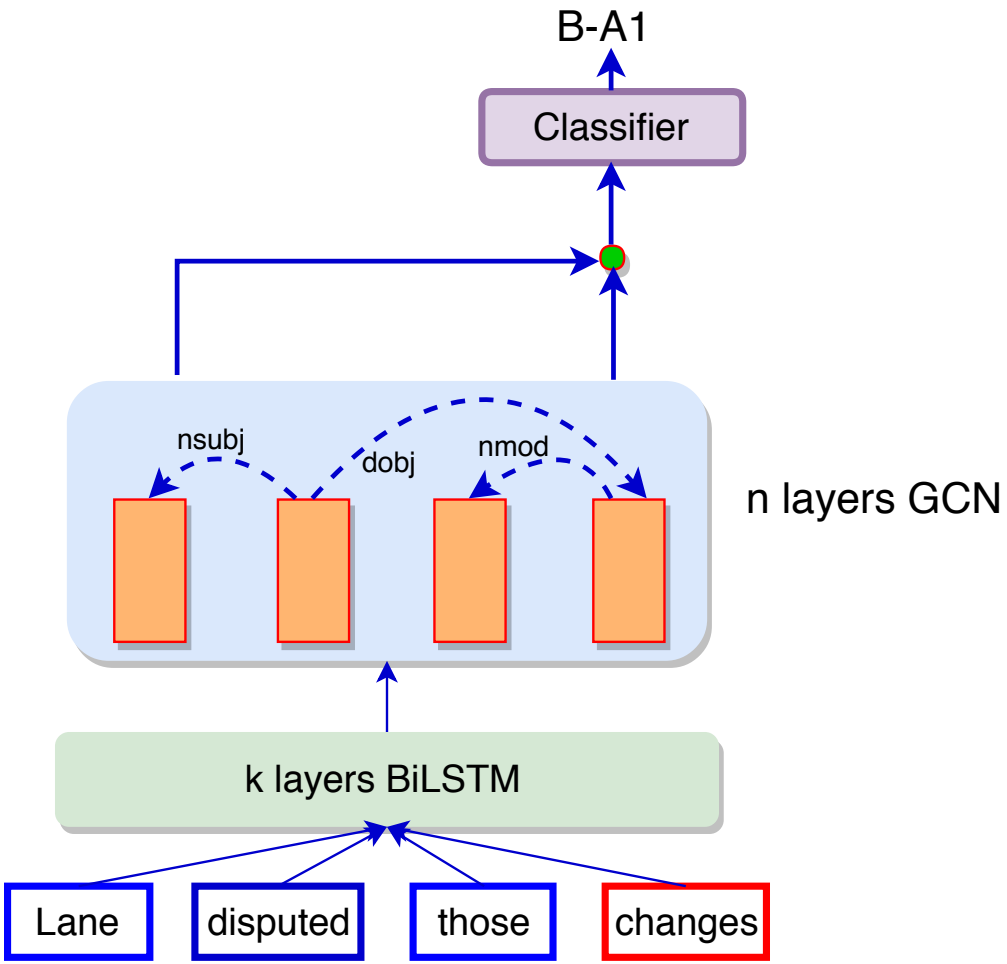


Figure: LSTM-GCN Model for Semantic Role Labeling (Marcheggiani and Titov, 2017)

References I

Björkelund, A., Bohnet, B., Hafdell, L., and Nugues, P. (2010). A high-performance syntactic and semantic dependency parser. In *Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations*, pages 33–36. Association for Computational Linguistics.

Gildea, D. and Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational linguistics*, 28(3):245–288.

He, L., Lee, K., Lewis, M., and Zettlemoyer, L. (2017). Deep semantic role labeling: What works and what's next. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 473–483.

Huang, Z., Xu, W., and Yu, K. (2015). Bidirectional lstm-crf models for sequence tagging. *arXiv preprint arXiv:1508.01991*.

Kingsbury, P. and Palmer, M. (2002). From treebank to propbank. In *LREC*, pages 1989–1993. Citeseer.

Kipf, T. N. and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Marcheggiani, D., Frolov, A., and Titov, I. (2017). A simple and accurate syntax-agnostic neural model for dependency-based semantic role labeling. *arXiv preprint arXiv:1701.02593*.

Marcheggiani, D., Roth, M., Titov, I., and Van Durme, B. (2013). Neural methods for semantic role labeling. *EMNLP 2017, Copenhagen*.

Marcheggiani, D. and Titov, I. (2017). Encoding sentences with graph convolutional networks for semantic role labeling. *arXiv preprint arXiv:1703.04826*.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.

Palmer, M., Wu, S., and Titov, I. (2013). Semantic role labeling tutorial. *NAACL, June 9, 2013*.

Pradhan, S., Hacioglu, K., Krugler, V., Ward, W., Martin, J. H., and Jurafsky, D. (2005). Support vector learning for semantic argument classification. *Machine Learning*, 60(1-3):11–39.

References II

Punyakanok, V., Roth, D., and Yih, W.-t. (2008). The importance of syntactic parsing and inference in semantic role labeling. *Computational Linguistics*, 34(2):257–287.

Täckström, O., Ganchev, K., and Das, D. (2015). Efficient inference and structured learning for semantic role labeling. *Transactions of the Association for Computational Linguistics*, 3:29–41.

Toutanova, K., Haghighi, A., and Manning, C. D. (2008). A global joint model for semantic role labeling. *Computational Linguistics*, 34(2):161–191.

Xue, N. and Palmer, M. (2004). Calibrating features for semantic role labeling. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*.

Zhou, J. and Xu, W. (2015). End-to-end learning of semantic role labeling using recurrent neural networks. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1127–1137.