Enhanced Universal Dependency Parsing with Second-Order Inference and Mixture of Training Data

Xinyu Wang, Yong Jiang, Kewei Tu

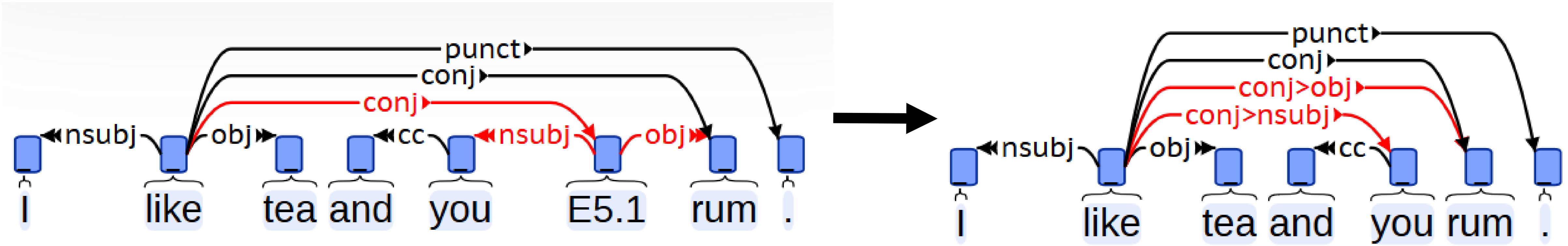
School of Information Science and Technology, ShanghaiTech University
DAMO Academy, Alibaba Group

Our Parser

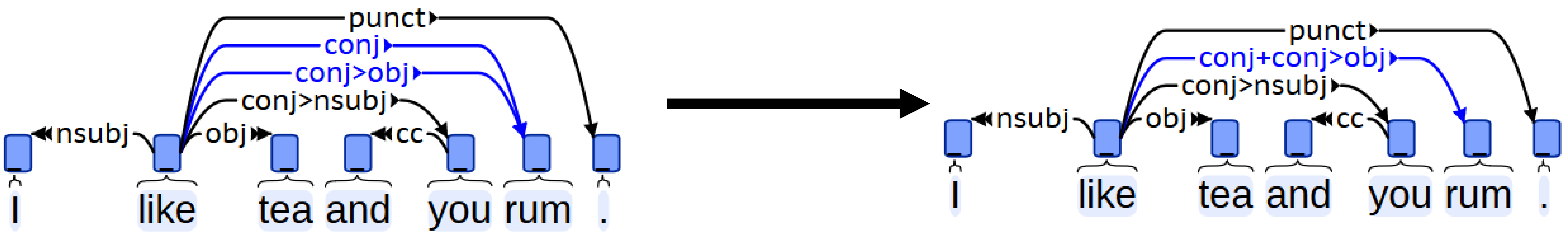
- A second-order semantic dependency parser based on Wang et al. (2019)
- Equip the parser with state-of-the-art contextual multilingual embeddings: XLM-R (Conneau et al., 2019)
- Improve the accuracy for the low-resource language (Tamil) through mixing the training set with another language (English/Czech)
- Our Parser performs 0.6 ELAS better than the best parser in official results after fixing the graph connectivity issues

[1]: Xinyu Wang, Jingxian Huang, and Kewei Tu. 2019. Second-order semantic dependency parsing with end-to-end neural networks.
[2]: Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Unsupervised cross-lingual representation learning at scale.

Preprocessing: Empty Nodes



Preprocessing: Repeated Edges

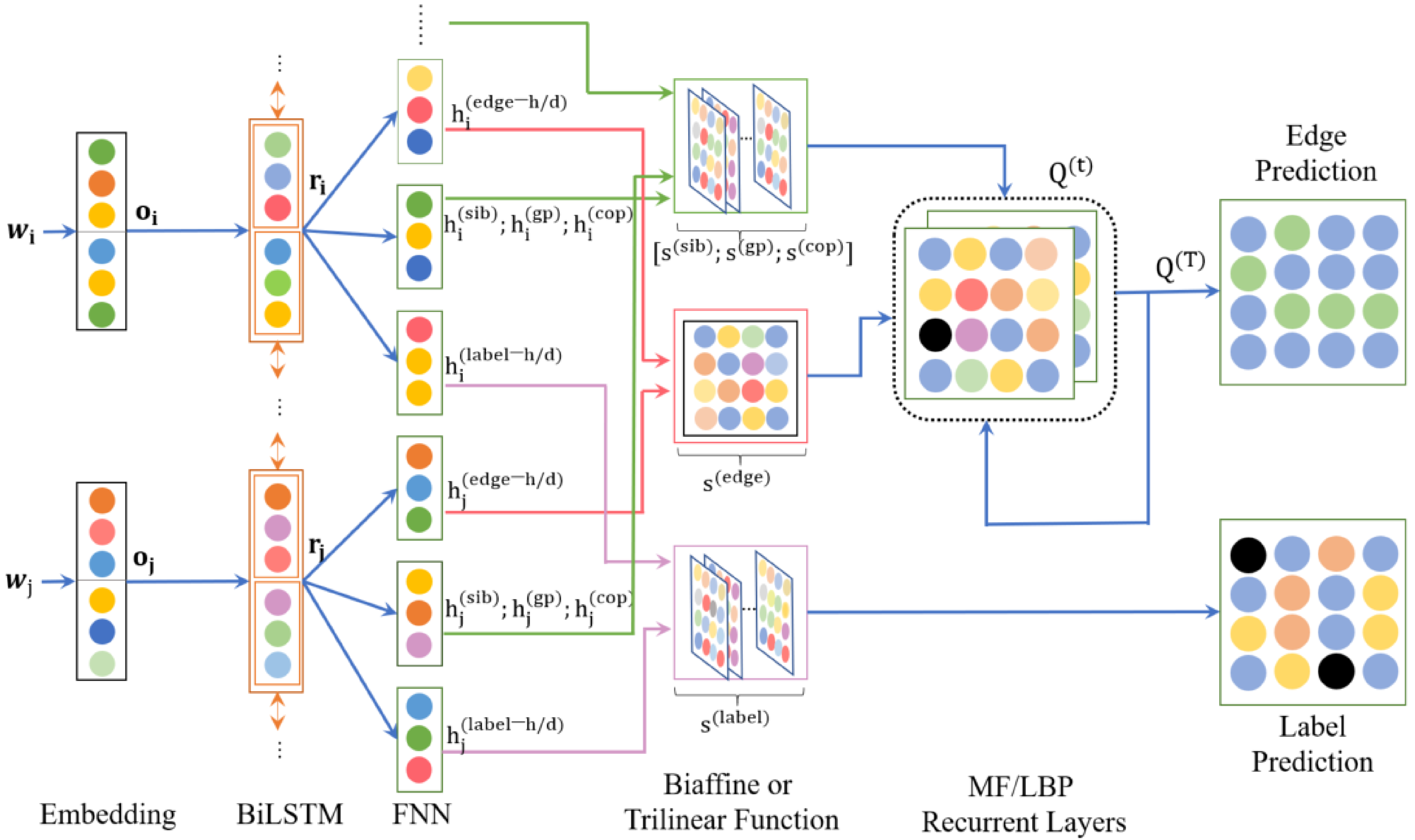


Preprocessing

- Tokenization: Stanza (Qi et al., 2020)
- Multiple Treebanks: concatenate the datasets
- Splitting the development sets into halves as validation and test sets

[1]: Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D Manning. 2020. Stanza: A python natural language processing toolkit for many human languages.

Approach (Wang et al., 2019)



Mixture of Training Data For Tamil

- Problem: low-resource
  - Only 400 training sentences for Tamil
- Solution: utilizing rich-resource language corpus
  - Multilingual Embedding: XLM-R
  - Rich-Resource languages: English (12k sents) or Czech (100k sents)
  - Remove the label of dependency edges in rich-resource training data
- New training data: Upsampled Tamil training data + rich-resource training data
- Additional language-specific embeddings: Flair (Akbik et al., 2018) and fastText (Bojanowski et al., 2017)

[1]: Alan Akbik, Duncan Blythe, and Roland Vollgraf.2018. Contextual string embeddings for sequence labeling.
[2]: Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information

Graph Connection

- Original submission:
  - Non-connected graphs (all potential edges with probability > 0.5)
- New solution:
  - Tree algorithms: Maximum Spanning Tree (MST) or Eisner's Algorithm
  - First use MST or Eisner's algorithm to keep connectivity of graphs and then add potential edges with probabilities larger than 0.5

Results

Team Name	ar	bg	cs	nl	en	et	fi	fr	it	lv	lt	pl	ru	sl	sv	ta	uk	Avg.
	Official																	
RobertNLP	0.0	0.0	0.0	0.0	**88.9**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.2
Koebsala	60.8	68.9	61.1	62.9	65.4	59.1	67.5	67.9	69.1	64.8	56.3	61.3	64.2	64.1	64.5	47.4	64.2	62.9
ADAPT	57.2	77.3	66.4	67.7	70.4	61.1	72.4	74.7	72.0	72.4	58.4	65.9	75.3	68.4	68.4	48.5	66.4	67.2
clasp	51.3	84.9	67.1	78.9	82.9	60.4	66.0	72.8	87.1	66.0	52.6	71.2	70.4	65.2	71.4	42.2	63.2	67.9
Ours	63.4	78.7	75.4	70.9	72.3	74.9	76.0	77.0	73.1	77.8	66.9	71.0	78.3	73.1	69.6	48.2	73.0	71.7
Unipi	57.8	84.9	76.0	77.6	84.0	57.2	72.1	78.9	89.1	68.2	61.1	70.6	76.9	81.4	78.7	48.5	73.9	72.8
FASTPARSE	66.9	84.9	77.2	77.4	78.5	74.1	75.7	77.8	84.8	75.6	61.4	74.5	80.4	73.5	75.2	47.0	74.0	74.0
EmoryNLP	67.3	88.2	85.5	80.7	85.3	81.4	83.0	**86.2**	88.5	79.2	66.1	82.4	88.6	82.7	78.2	54.3	79.7	79.8
OrangeDeskin	71.0	89.4	87.0	85.1	85.2	81.0	86.2	83.6	90.8	82.1	75.9	80.4	89.8	84.4	83.3	**64.2**	84.6	82.6
TurkuNLP	**77.8**	90.7	87.5	84.7	87.2	84.5	**89.5**	85.9	91.5	84.9	77.6	**84.6**	90.7	**88.6**	**85.6**	57.8	87.2	84.5
	Post-Evaluation																	
Ours+en+MST	77.7	91.5	90.1	86.2	87.1	86.0	89.0	85.3	91.5	87.6	78.9	84.0	92.3	87.6	84.7	56.7	88.0	85.0
Ours+cs+Eis	77.8	91.1	89.5	86.3	87.2	85.7	88.5	85.3	**91.5**	87.3	78.6	83.7	92.3	87.1	84.8	58.4	88.0	84.9
Ours+cs+MST	77.7	**91.5**	**90.1**	**86.2**	87.1	**86.0**	89.0	85.3	91.5	**87.6**	**78.9**	84.0	**92.3**	87.6	84.7	58.4	**88.0**	**85.1**

Results

Team Name	ar	bg	cs	nl	en	et	fi	fr	it	lv	lt	pl	ru	sl	sv	ta	uk	Avg.
									Official									
RobertNLP	0.0	0.0	0.0	0.0	**88.9**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.2
Koebsala	60.8	68.9	61.1	62.9	65.4	59.1	67.5	67.9	69.1	64.8	56.3	61.3	64.2	64.1	64.5	47.4	64.2	62.9
ADAPT	57.2	77.3	66.4	67.7	70.4	61.1	72.4	74.7	72.0	72.4	58.4	65.9	75.3	68.4	68.4	48.5	66.4	67.2
clasp	51.3	84.9	67.1	78.9	82.9	60.4	66.0	72.8	87.1	66.0	52.6	71.2	70.4	65.2	71.4	42.2	63.2	67.9
Ours	63.4	78.7	75.4	70.9	72.3	74.9	76.0	77.0	73.1	77.8	66.9	71.0	78.3	73.1	69.6	48.2	73.0	71.7
Unipi	57.8	84.9	76.0	77.6	84.0	57.2	72.1	78.9	89.1	68.2	61.1	70.6	76.9	81.4	78.7	48.5	73.9	72.8
FASTPARSE	66.9	84.9	77.2	77.4	78.5	74.1	75.7	77.8	84.8	75.6	61.4	74.5	80.4	73.5	75.2	47.0	74.0	74.0
EmoryNLP	67.3	88.2	85.5	80.7	85.3	81.4	83.0	**86.2**	88.5	79.2	66.1	82.4	88.6	82.7	78.2	54.3	79.7	79.8
OrangeDeskin	71.0	89.4	87.0	85.1	85.2	81.0	86.2	83.6	90.8	82.1	75.9	80.4	89.8	84.4	83.3	**64.2**	84.6	82.6
TurkuNLP	**77.8**	90.7	87.5	84.7	87.2	84.5	**89.5**	85.9	91.5	84.9	77.6	**84.6**	90.7	**88.6**	**85.6**	57.8	87.2	84.5
									Post-Evaluation									
Ours+en+MST	77.7	91.5	90.1	86.2	87.1	86.0	89.0	85.3	91.5	87.6	78.9	84.0	92.3	87.6	84.7	56.7	88.0	85.0
Ours+cs+Eis	77.8	91.1	89.5	86.3	87.2	85.7	88.5	85.3	**91.5**	87.3	78.6	83.7	92.3	87.1	84.8	58.4	88.0	84.9
Ours+cs+MST	77.7	**91.5**	**90.1**	**86.2**	87.1	**86.0**	89.0	85.3	91.5	**87.6**	**78.9**	84.0	**92.3**	87.6	84.7	58.4	**88.0**	**85.1**

Mixture of Data Comparison

Combination	# Training Sentences	ELAS
Tamil	400	55.39
English+Tamil	12543+12400	56.56
Czech+Tamil	102131+102000	**58.44**

First-Order vs. Second-Order and Concatenating Other Embeddings

Approach	ar	bg	cs	nl	en	et	fi	fr	it
XLMR+Flair+FastText+1st-Order	81.66	89.29	91.04	92.55	89.74	88.33	89.40	90.64	91.94
XLMR+Flair+FastText+2nd-Order	81.98	89.43	**91.39**	**92.68**	89.58	**88.69**	89.54	91.08	**91.98**
XLMR+1st-Order	82.02	90.15	90.80	92.43	90.05	88.13	89.51	91.14	91.96
XLMR+2nd-Order	**82.42**	**90.37**	91.21	92.66	**90.26**	88.60	90.35	**91.69**	**91.98**
	lv	lt	pl	ru	sk	sv	ta	uk	Avg.
XLMR+Flair+FastText+1st-Order	88.21	80.21	86.91	92.88	87.28	85.52	66.17	88.26	89.40
XLMR+Flair+FastText+2nd-Order	88.59	81.25	86.46	**93.28**	87.18	85.63	**68.76**	88.04	89.59
XLMR+1st-Order	89.62	81.92	85.73	92.86	88.48	86.36	63.28	88.96	89.57
XLMR+2nd-Order	**89.97**	**83.24**	**87.49**	93.21	**89.07**	**86.85**	64.84	**89.99**	**89.95**

*: We use labeled F1 score here, which is the metric for SDP

Comparisons of Graph Connection Approaches (Treebank Level)

Graph	ar-PADT	bg-BTB	cs-FicTree	cs-CAC	cs-PDT	cs-PUD	nl-Alpino	nl-LassySmall
Non-Connected	77.74	**91.50**	**90.60**	90.55	**90.65**	**84.26**	**90.11**	82.55
MST	77.73	91.48	90.51	**90.59**	90.63	84.25	90.09	82.51
Eisner's	**77.75**	91.07	89.85	90.02	90.05	83.70	89.69	**83.10**

Graph	en-EWT	en-PUD	et-EDT	et-EWT	fi-TDT	fi-PUD	fr-Sequoia	fr-FQB
Non-Connected	86.33	**88.05**	**87.36**	**79.62**	**90.00**	**87.52**	89.67	84.11
MST	86.30	**88.05**	87.34	79.61	89.97	**87.52**	89.66	84.09
Eisner's	**86.40**	88.04	87.07	79.42	89.44	86.97	**89.73**	**84.12**

Graph	it-ISDT	lv-LVTB	lt-ALKSNIS	pl-LFG	pl-PDB	pl-PUD	ru-SynTagRus	sl-SNK
Non-Connected	91.50	**87.69**	78.97	**87.65**	**83.23**	**82.96**	**92.62**	**87.56**
MST	91.49	87.64	**78.94**	**87.65**	83.21	82.95	92.31	87.55
Eisner's	**91.52**	87.29	78.63	87.59	82.90	82.57	92.31	87.14

Graph	sv-Talbanken	sv-PUD	ta-TTB	uk-IU	Average
Non-Connected	88.35	**80.88**	56.51	88.00	**85.59**
MST	88.33	80.87	56.56	**88.02**	85.57
Eisner's	**88.37**	80.87	**56.71**	**88.02**	85.37

Comparisons of Graph Connection Approaches (Language Level)

Team Name	ar	bg	cs	nl	en	et	fi	fr	it
TurkuNLP	**77.82**	90.73	87.51	84.73	**87.15**	84.54	**89.49**	**85.90**	**91.54**
Ours+en+MST	77.74	91.48	90.09	86.19	87.10	85.97	88.99	85.28	91.49
Ours+en	77.75	**91.50**	**90.11**	**86.22**	87.12	**85.99**	89.01	85.29	91.50

Team Name	lv	lt	pl	ru	sl	sv	ta	uk	Avg.
TurkuNLP	84.94	77.64	**84.64**	90.69	**88.56**	**85.64**	**57.83**	87.22	84.50
Ours+en+MST	87.64	78.94	84.00	92.31	87.55	84.74	56.71	**88.02**	84.96
Ours+en	**87.69**	**78.97**	84.01	**92.62**	87.56	84.75	56.52	88.00	**84.98**

Thank you

- Paper: https://arxiv.org/abs/2006.01414