

Why Attention? Analyze BiLSTM Deficiency and Its Remedies in the Case of NER

Peng-Hsuan Li

Tsu-Jui Fu

Wei-Yun Ma

jacobvsdanniel@iis.sinica.edu.tw

tsu-juifu@ucsb.edu

ma@iis.sinica.edu.tw

XOR Limitation for BiLSTM

- 1. Key and Peele (*work-of-art*)
- 2. You and I (work-of-art)
- 3. Key and I
- 4. You and Peele

$$s_{t} = H_{t}W_{p} + b$$

$$= \overrightarrow{H}_{t}\overrightarrow{W}_{p} + \overleftarrow{H}_{t}\overleftarrow{W}_{p} + b$$

$$= \overrightarrow{s}_{t} + \overleftarrow{s}_{t} + b$$

i: inside-work-of-art

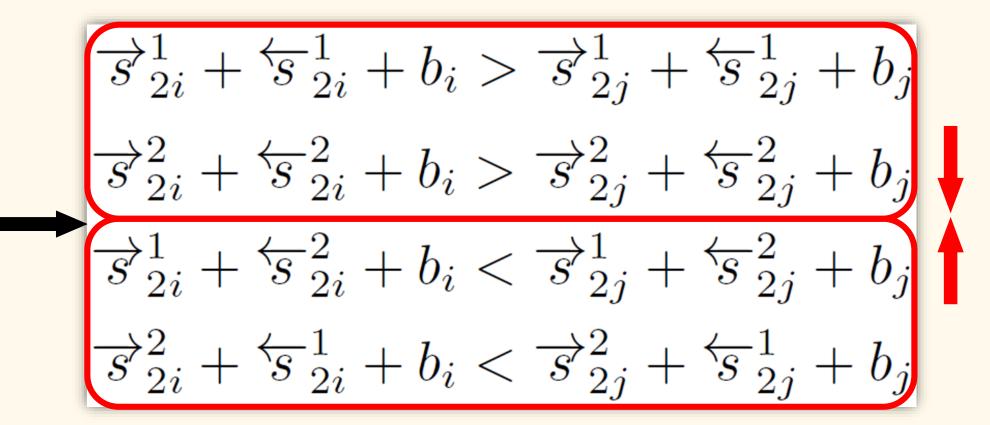
j: outside

$$\overrightarrow{s}_{2i}^{1} + \overleftarrow{s}_{2i}^{1} + b_{i} > \overrightarrow{s}_{2j}^{1} + \overleftarrow{s}_{2j}^{1} + b_{j}$$

$$\overrightarrow{s}_{2i}^{2} + \overleftarrow{s}_{2i}^{2} + b_{i} > \overrightarrow{s}_{2j}^{2} + \overleftarrow{s}_{2j}^{2} + b_{j}$$

$$\overrightarrow{s}_{2i}^{3} + \overleftarrow{s}_{2i}^{3} + b_{i} < \overrightarrow{s}_{2j}^{3} + \overleftarrow{s}_{2j}^{3} + b_{j}$$

$$\overrightarrow{s}_{2i}^{4} + \overleftarrow{s}_{2i}^{4} + b_{i} < \overrightarrow{s}_{2j}^{4} + \overleftarrow{s}_{2j}^{4} + b_{j}$$



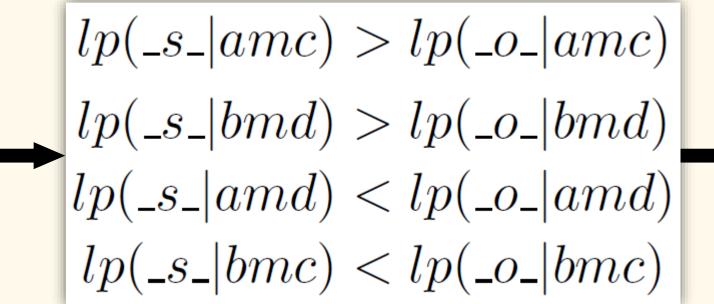
XOR Limitation for CRF-BiLSTM

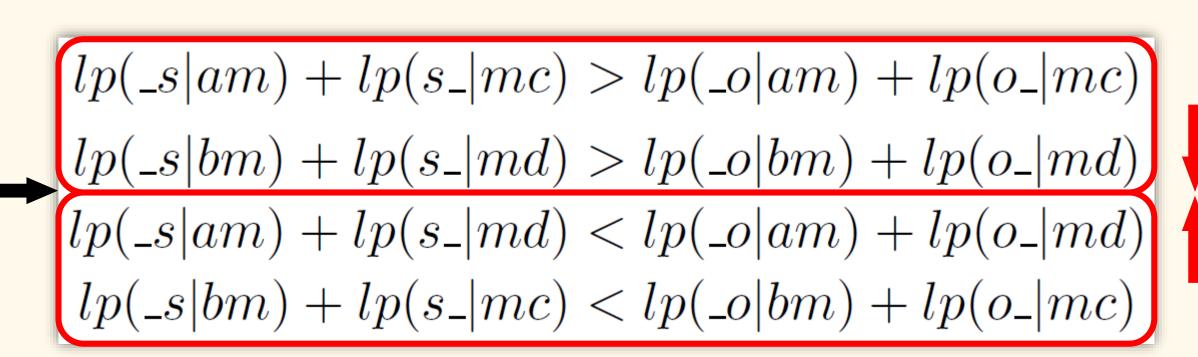
- 1. $a_{-}o m_{-}s c_{-}o$
- 2. $b_{-}o \ m_{-}s \ d_{-}o$

3. $a_{-}o \ m_{-}o \ d_{-}o$

- Viterbi with softmax (p) and transition (q)
- 4. $b_{-}o m_{-}o c_{-}o$ Viterbi, suppose no transition

p(oso|amc) > p(ooo|amc)p(oso|bmd) > p(ooo|bmd)p(oso|amd) < p(ooo|amd)p(oso|bmc) < p(ooo|bmc)





 $lp(_s|am) + lp(s_|mc) + lq(oso) > lp(_o|am) + lp(o_|mc) + lq(ooo)$

 $lp(_s|bm) + lp(s_|md) + lq(oso) > lp(_o|bm) + lp(o_|md) + lq(ooo)$

 $lp(_s|bm) + lp(s_|mc) + lq(oso) < lp(_o|bm) + lp(o_|mc) + lq(ooo)$

 $lp(_s|am) + lp(s_|md) + lq(oso) < lp(_o|am) + lp(o_|md) + lq(oso)$

BILSTM $\overrightarrow{H} = \overrightarrow{LSTM}_2(\overrightarrow{LSTM}_1(X))$

$$\overleftarrow{H} = \overleftarrow{LSTM}_4(\overleftarrow{LSTM}_3(X))$$

$$H = \overrightarrow{H} \mid | \overleftarrow{H},$$

$$s_t = H_t W_p + b$$

-> Cross

$$\overrightarrow{H}^{1} = \overrightarrow{LSTM}_{1}(X)$$

$$\overleftarrow{H}^{3} = \overleftarrow{LSTM}_{3}(X)$$

$$\overrightarrow{H}^{2} = \overrightarrow{LSTM}_{2}(\overrightarrow{H}^{1}||\overleftarrow{H}^{3})$$

$$\overleftarrow{H}^{4} = \overleftarrow{LSTM}_{4}(\overrightarrow{H}^{1}||\overleftarrow{H}^{3})$$

$$H = \overrightarrow{H}^{2}||\overleftarrow{H}^{4}$$

-> **Att**

$$\alpha^{i} = \sigma \left(\frac{HW^{qi}(HW^{ki})^{T}}{\sqrt{d_{c}}} \right)$$

$$C^{i} = \alpha^{i}HW^{vi},$$

$$s_{t}^{c} = (H_{t}||C_{t}^{1}||C_{t}^{2}||...||C_{t}^{m})W_{c} + b$$

Results

		OntoNo	tes 5.0	WNUT 2017				
	Prec.	Rec.	F1	Prec.	Rec.	F1		
BiLSTM-CNN	86.04	86.53	86.28 ± 0.26	_	_	_		
CRF-IDCNN	-	-	86.84 ± 0.19	-	-	_		
CRF-BiLSTM(-BiLSTM*)	-	-	86.99 ± 0.22	-	-	38.24		
Baseline-BiLSTM-CNN	88.37	87.14	87.75 ± 0.14	53.24	32.93	40.68 ± 1.78		
Cross-BiLSTM-CNN	88.37	88.17	88.27 ± 0.17	58.28	33.92	42.85 ± 0.99		
Att-BiLSTM-CNN	88.71	88.11	88.40 ± 0.18	55.82	34.08	42.26 ± 0.82		

-> By Length

		OntoNo	otes 5.0		WNUT 2017				
	1	2	3	3+	1	2	3	3+	
Cross	+0.3%	+0.6%	+1.8%	+1.3%	+1.7%	+2.9%	+8.7%	+5.4%	
Att	+0.1%	+1.1%	+2.3%	+1.8%	+1.5%	+2.0%	+2.6%	+0.9%	

-> By Component

	Baseline								
	HC^{all}	H	C^{all}	C^1	C^2	C^3	C^4	C^5	H
O	99.05	-1.68	0.75	0.95	-1.67	-45.57	-0.81	-35.46	-0.03
S	93.74	2.69	-91.02	-90.56	-90.88	-25.61	-86.25	-84.32	0.13
В	90.99	1.21	-52.26	-90.78	-88.08	-90.88	-12.21	-87.45	-0.63
I	90.09	-28.18	-3.80	-87.93	-60.56	-50.19	-57.19	-79.63	-0.41
Е	93.23	2.00	-71.50	-93.12	-36.45	-39.19	-91.90	-90.83	-0.38

Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	a r	neeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the		house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White		aturda	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Teraali

							The state of the s				
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	a _I	neeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the		house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White		aturda	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Israeli
Clinton	will	host	а	meeting	at	the	White	house	Saturday	with	Igraali