A Span Selection Model for Semantic Role Labeling

Hiroyuki Shindo

Yuji Matsumoto



Center for **Advanced Intelligence Project** hiroki.ouchi@riken.jp, {shindo, matsu}@is.naist.jp



Hiroki Ouchi



Span Selection Model

Span Selection SRL Model

Code: https://github.com/hiroki13/span-based-srl

English Semantic Role Labeling Task



She <u>bought</u> a small dog last summer **A0** A1 TMP

Challenge

• How to model spans (phrases) in a sentence?

Solution

- Solve SRL as a span selection problem
- Model all possible spans in a sentence

Contribution

- A simple span-based model that achieves the SOTA (about 87.0 F1)
- Quantitative and qualitative analysis on our span-based model
- Empirical analysis on the performance gains by ELMo

Procedure

- (1) List all possible spans in an input sentence
- (2) Score each span
- (3) Select higher scoring spans





Overall Results

			CoNLL-05		CoNLL-12
Emb	Model	WSJ	Brown	ALL	
	CRF	82.9	71.3	81.4	82.4
SENNA	Span	83.5	73.1	82.1	83.0
	Span(Ens)	85.1	74.8	83.7	84.7
	CRF	87.3	78.4	86.1	85.9
ELMo	Span	87.6	78.7	86.4	86.2
	Span(Ens)	88.5	79.6	87.4	87.0

- Our span-based model consistently outperforms the CRF model \bullet
- ELMo improves the performance by a large margin

Comparison with Existing Models

	CoNLL-05	CoNLL-12	
WSJ	Brown	ALL	

Span Boundary Match & Label Prediction

		CoNLL-05		CoNLL-12	
Emb	Model	Span-F1	Label-Acc.	Span-F1	Label-Acc.
SENNA	Span	86.6	95.3	87.3	95.1
	CRF	87.0	93.8	87.9	93.6
ELMo	Span	90.5	96.1	90.3	95.7
	CRF	91.2	95.2	90.9	94.4

- CRF > Span Span Boundary Match: lacksquare
- Label Prediction: CRF < Span \bullet
- ELMo improves the performance for span boundary match

Analysis on Span Representations

`` ... toy makers to move [across the border] "

GOLD : A2

PRED : DIR

Single Model					
ELMo-Span	87.6	78.7	86.4	86.2	
He+ 18	87.4	80.4	-	85.5	
Tan+ 18	84.8	74.1	83.4	83.4	
He+ 17	83.1	72.1	81.6	81.7	
Zhou+ 15	82.8	69.4	81.1	81.3	
Ensemble Model					
ELMo-Span	88.5	79.6	87.4	87.0	
Tan+ 18	86.1	74.8	84.6	83.9	
He+ 17	84.6	73.6	83.2	83.4	
FitzGerald+ 15	80.3	72.2	-	80.1	

Our span-based model achieves the state-of-the-art results

	Near	rest neighbors of ``across the border"
1	DIR	across the Hudson
2	DIR	outside their traditional tony circle
3	DIR	across the floor
4	DIR	through this congress
5	A2	off their foundations
6	DIR	off its foundation
7	DIR	off the center field wall
8	A3	out of bed
9	A2	through cottage rooftops
10	DIR	through San Francisco

Spans with a misclassified label often have their nearest neighbors with inconsistent labels