CS224n HW2

Xinyu Tan

August 2, 2017

1 Neural Transition-based Dependency Parsing

(a)

stack	buffer	new dependency	transition
[root]	[I, parsed, this, sentence, correctly]		Initial Configuratio
[root, I]	[parsed, this, sentence, correctly]		SHIFT
[root, I, parsed]	[this, sentence, correctly]		SHIFT
[root, parsed]	[this, sentence, correctly]	parsed $\rightarrow$ I	LEFT-ARC
[root, parsed, this]	[sentence, correctly]		SHIFT
[root, parsed, this, sentence]	[correctly]		SHIFT
[root, parsed, sentence]	[correctly]	sentence $\rightarrow$ this	LEFT-ARC
[root, parsed]	[correctly]	parsed $\rightarrow$ sentence	RIGHT-ARC
[root, parsed, correctly]	[]		SHIFT
[root, parsed]	[]	parsed $\rightarrow$ correctly	RIGHT-ARC
[root]	[]	root $\rightarrow$ parsed	RIGHT-ARC

(b)

The sentence will be parsed in $2n$ times. Each word will be pushed into stack once, and each word only depends on one other word. Therefore, the process is in $O(n)$ time complexity.

(f)

We need to satisfy: $\mathbb{E}_{p_{\text{drop}}}[\boldsymbol{h}_{\text{drop}}]_i = \gamma(1 - p_{\text{drop}})\boldsymbol{h}_i = \boldsymbol{h}_i$, then we have:

$$\gamma = \frac{1}{1 - p_{\text{drop}}}$$