

1- (a)

i.

가속의 momentum과 달리 ~~각시점평균~~ (rolling average)를  
사용하기 때문에  $\mu$ 의 계수가 작은 값을 가진다.

이때 따라 뒤로 갈수록 최신 계수가 전체 평균에  $(1-\beta)^n$ 만큼의  
영향을 끼친다. 따라서 현재  $(10)$ 이거  
noise 혹은 이런 train data 이라는 것을 보면,  
해당 gradients의 많은 영향을 받기 좋을 것야...  
런칭이니까 점점 점점 바뀌는 거야.

$\Rightarrow$  low variance를 갖게 될것이야.

$\Rightarrow$  즉, 특정 data에 민감하게 반응하기 않으므로  
overfitting이 줄어들 것이다.

i)

m은 매번 바뀌는 것임이 아니라 값이 늘었다, 줄었다.

r는 제곱이었 제곱 커짐!

$\frac{m}{\sqrt{r}}$  : 학습 n step이 너무 크면 학습 진행하는 방향

1-(b)

$$x' = x + \epsilon$$

$$E(x') = x$$

i)

$$h' = \begin{cases} 0 & , p_{drop} \\ \frac{h}{1-p} & , \text{그 외} \end{cases}$$

$$h_i = [p_{drop} \cdot 0 + (1-p_{drop}) \cdot 1] r$$

$$\therefore r = \frac{1}{1-p_{drop}}$$

ii)

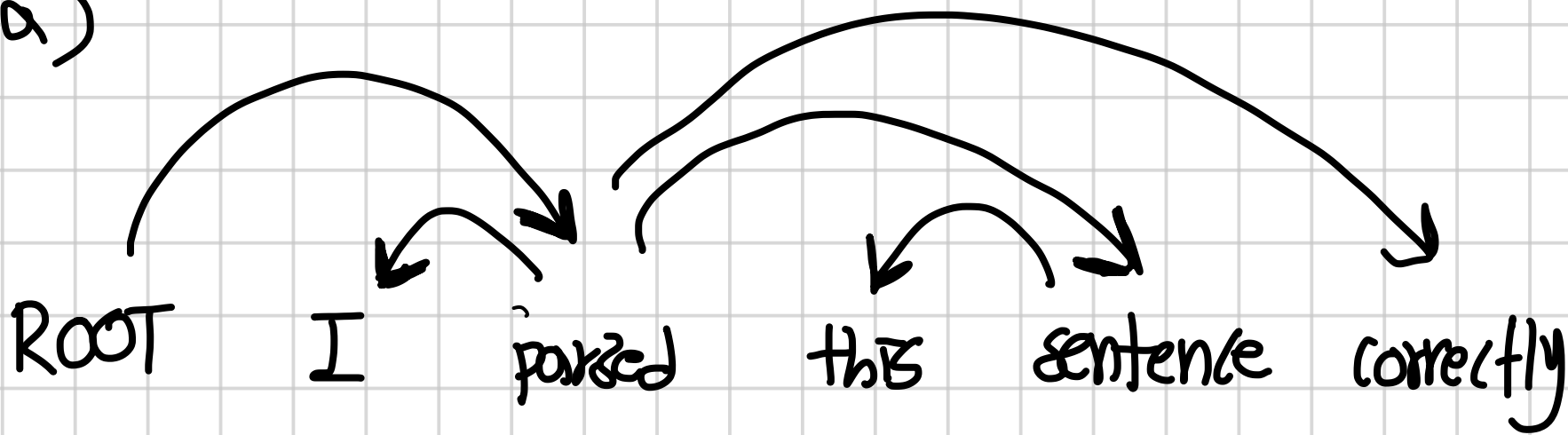
① dropout은 의도적으로 노이즈를 주는 것임 때문에

'잘못된' 값이 나옴

② random하게 발생하여 때문에

같은 input에 따라 값이 달라짐!

2-(a)



Stack	Buffer	New dependency	Transitions
[ROOT]	[I, parsed, this, sentence, correctly]		(init)
[Root, I]	[parsed, this, sentence, correctly]		
[Root, I, <del>parsed</del> ]	[this, sentence, correctly]		
[Root, <del>parsed</del> ]	[this, sentence, correctly]	parsed → I	
[Root, <del>parsed</del> , this]	[sentence, correctly]		
[Root, <del>parsed</del> , this, sentence]	[correctly]		
[Root, <del>parsed</del> , sentence]		parsed → I, sentence → this	
[Root, <del>parsed</del> ]		<del>parsed</del> → I, parsed → sentence	
[Root, <del>parsed</del> , correctly]		parsed → I, parsed → sentence, parsed → correctly	
[Root, <del>parsed</del> ]			
[Root]		parsed → I, parsed → sentence, <del>parsed</del> → correctly, Root → parsed	

2-(b)

N

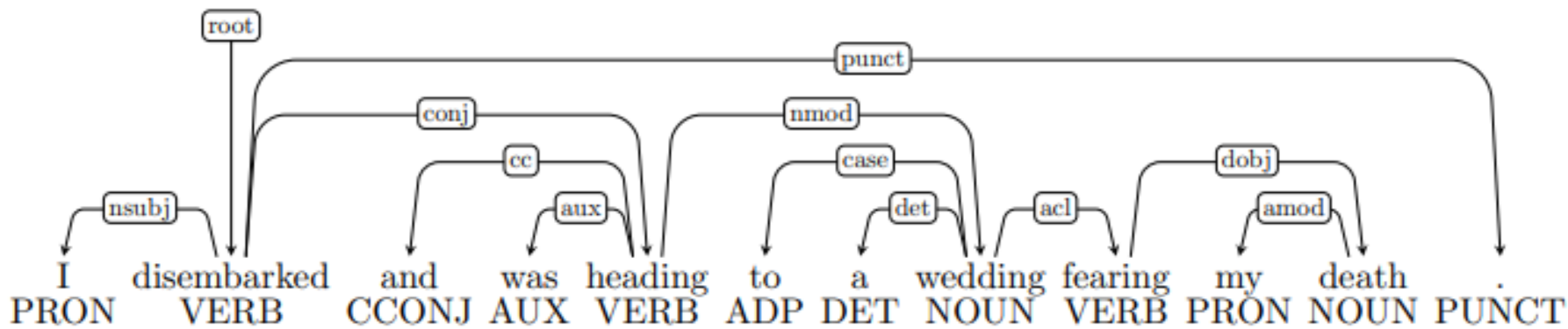
(?)

(step 1 transition 4/21)

2u

3-(f)

i.

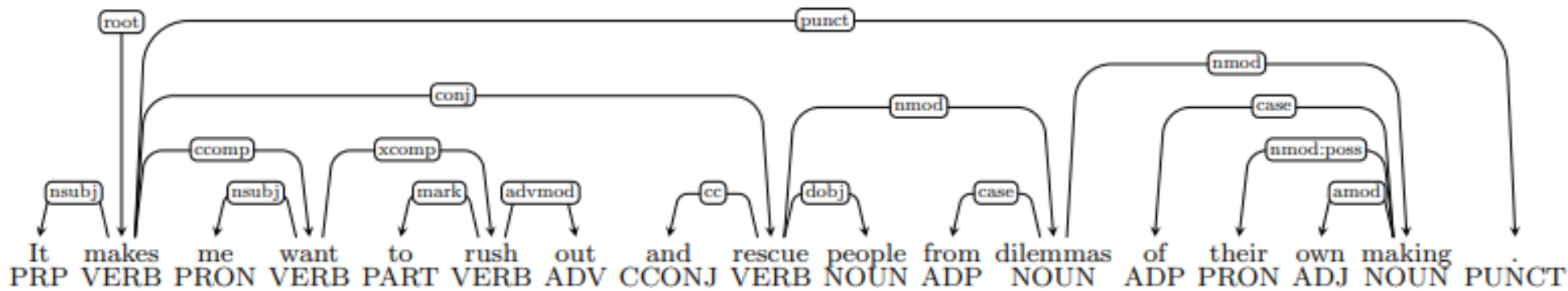


Error type :

Incorrect dependency :

Correct dependency :

ii.

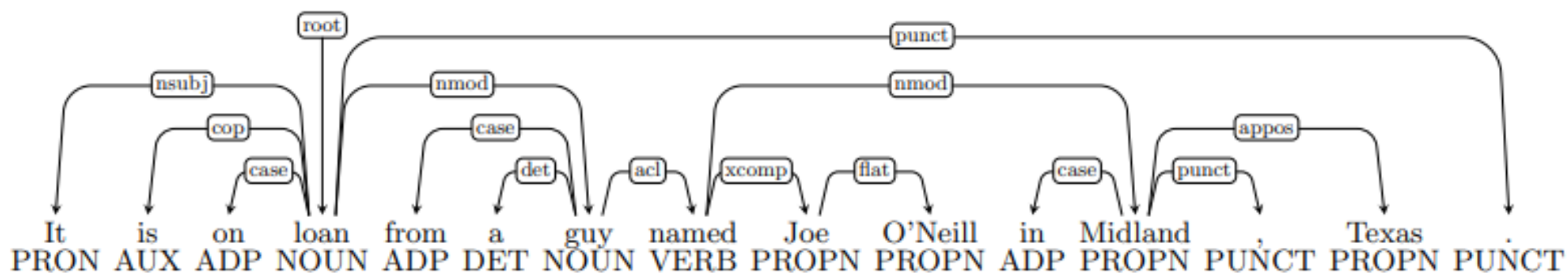


Error type: Prepositional Phrase Attachment Error

Incorrect dependency: makes → rescue

Correct dependency: want → rescue

iii.

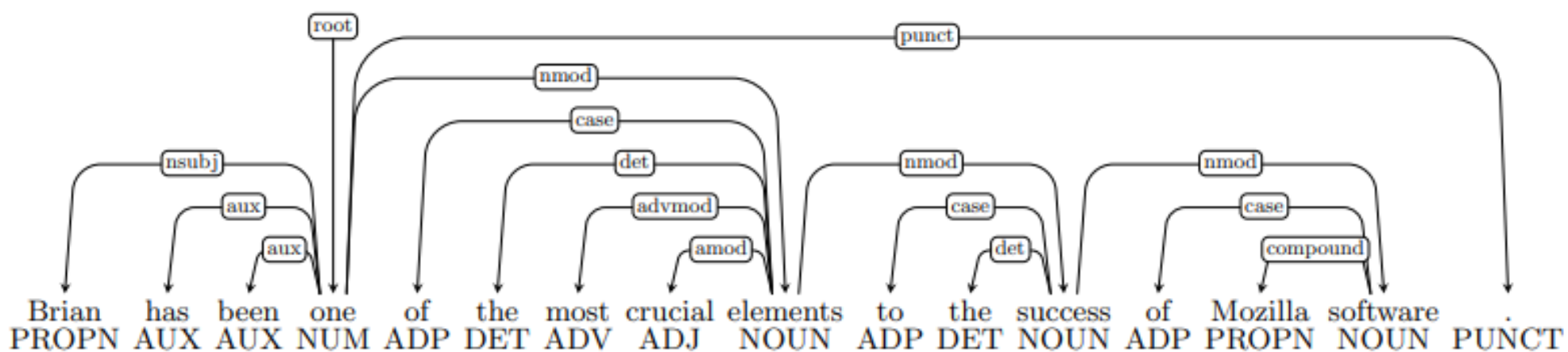


Error type: Prepositional Phrase

Incorrect dependency: Named → Midland

Correct dependency: guy → Midland

iv.



Error type: Verb phrase attachment error

Incorrect dependency: root → one

Correct dependency: Root → been