

(1)

(a)

(i)

M keeps track of the gradients before, and the general direction it is heading. That way each gradient is going to be less noisy, and it will still trend in the direction that was determined before.

(ii)

v is kept as a sort of normalization factor. When the m is close to zero, the process needs to be "sped up" So it is divided by a number that is also very small. If m has a large absolute value, it is "slowed down" to prevent overshooting, and divided by a big number.

(b)

(i)

Since the expected value of $h_{dropout}$ has to equal to h_i

$$\gamma(0 * p_{drop} + 1 * (1 - p_{drop}))h_i = h_i\gamma(1 - p_{drop})h_i = h_i\gamma(1 - p_{drop}) = 1\gamma = 1/(1 - p_{drop})$$

(ii)

The point of dropout is to make sure that the neurons do not rely too much on each other to make descisions. So once trained with dropouts, the neurons are able to make descisions independently, so there is no need for dropout.

(2)

(a)

Stack	Buffer	New Dependency	Transition
[ROOT]	[I, parsed, this, sentence, correctly]		Initial Configuration
[ROOT, I]	[parsed, this, sentence, correctly]		SHIFT
[ROOT, I, parsed]	[this, sentence, correctly]		SHIFT
[ROOT, parsed]	[this, sentence, correctly]	parsed → I	LEFT-ARC
[ROOT, parsed, this]	[sentence, correctly]		SHIFT
[ROOT, parsed, this, sentence]	[correctly]		SHIFT
[ROOT, parsed, sentence]	[correctly]	sentence → this	LEFT-ARC
[ROOT, parsed]	[correctly]	parsed → sentence	RIGHT-ARC
[ROOT, parsed, correctly]	□		SHIFT
[ROOT, parsed]	□	parsed → correctly	RIGHT-ARC
[ROOT]	□	ROOT → parsed	RIGHT-ARC

(b)

It will be parsed in $2N$ time. Each word has to be shifted, and then they should be a dependent. So there are two operations each word.

(f)

1.

Error type: Verb Phrase

Incorrect Dependency: Wedding \rightarrow fearing

Correct Dependency: heading \rightarrow fearing

2.

Error type: Coordination Attachment

Incorrect Dependency: makes \rightarrow rescue

Correct Dependency: rush \rightarrow rescue

3.

Error type: Prepositional Phrase

Incorrect Dependency: named \rightarrow Midland

Correct Dependency: guy \rightarrow Midland

4.

Error type: Modifier

Incorrect Dependency: elements \rightarrow most

Correct Dependency: crucial \rightarrow most