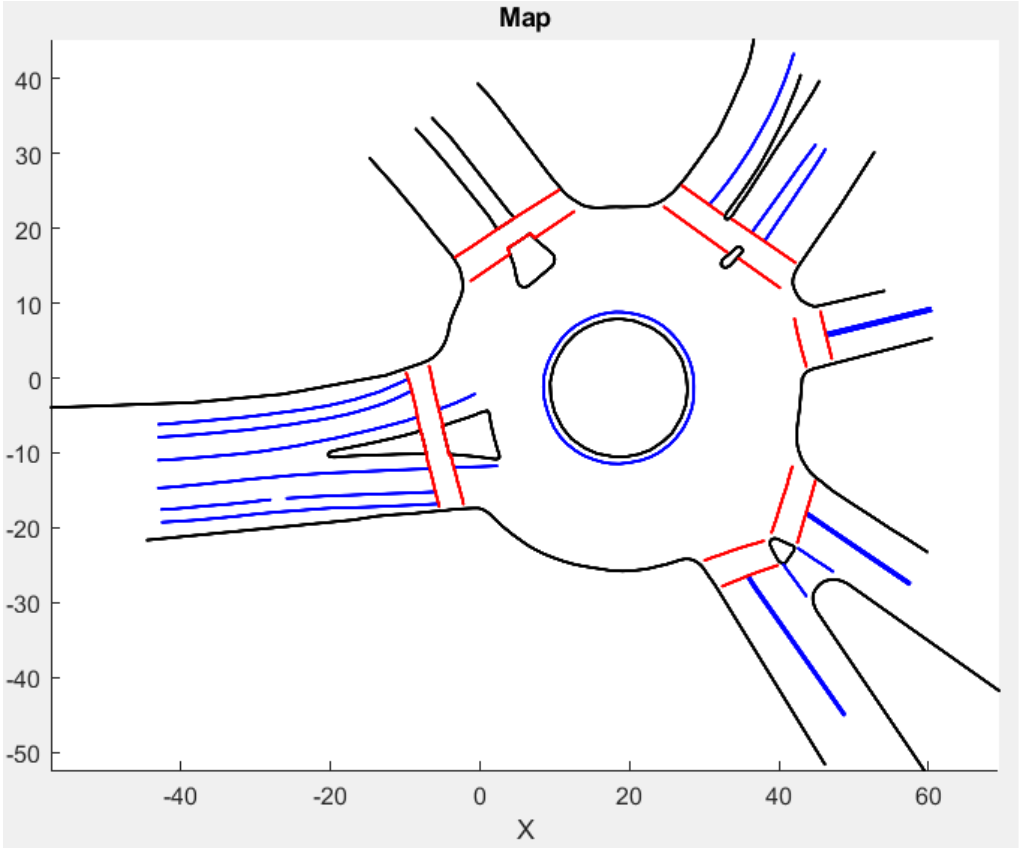
Vehicle Interaction Learning

Xiaosong Jia

08/21/2019

USA_Roundabout_FT



- Stop sign: <1s -> No interaction, >3s -> Exist interaction

Passing car's trajectory distance from stopping car less than 30 meters until 0

- TTC: <3s -> Exist interaction, >8s No interaction

Both cars' trajectory distance less than 30 meters from collision point until one of the cars pass that point

- No TTC: No interaction

- Remove samples with interaction time >20s

	Positive	Negative
Stop Sign	3185	377
TTC	1669	515
Total	4954	892

USA_Roundabout_FT

- Downsample: fps=5
- Clip: max_length = 20s

	Sample_len_avg	Label_len_avg
HighD	57.0 frames	14.5 frames
NGSIM	80.2 frames	14.8 frames
FT	57.4 frames	25.6 frames

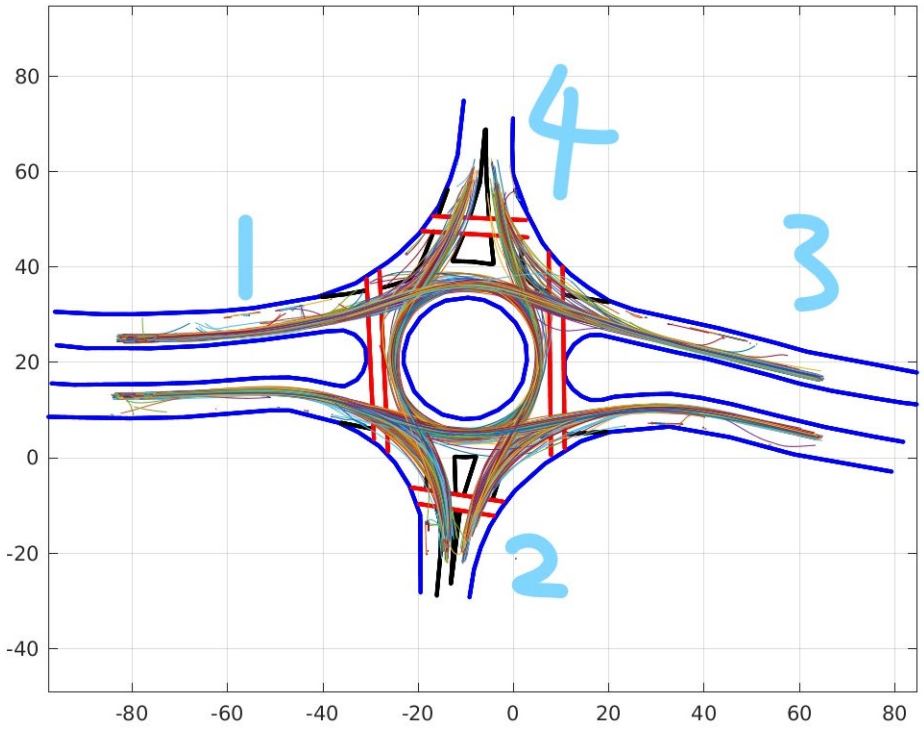
	IoU0.6 Acc	IoU0.9 Acc	Traj Cls Acc
HighD	94.8%	32.5%	99.1%
NGSIM	82.5%	18.9%	89.2%
FT	82.3%	35.1%	86.3%

USA_Roundabout_SR

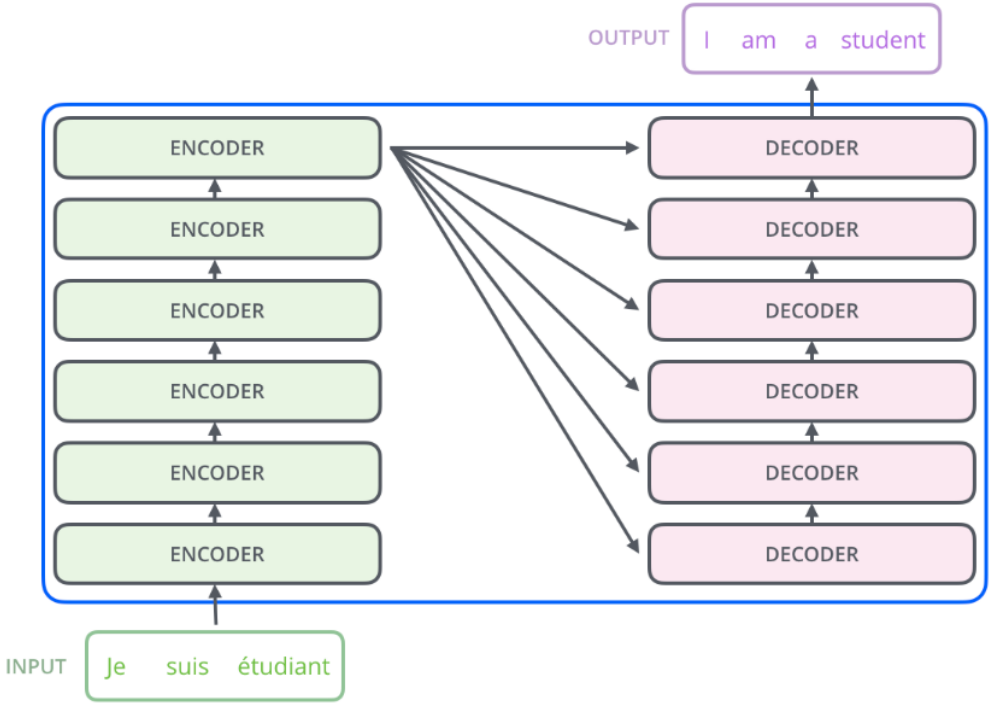
	Positive	Negative
Stop sign	22	2
TTC	213	124
Total	235	126

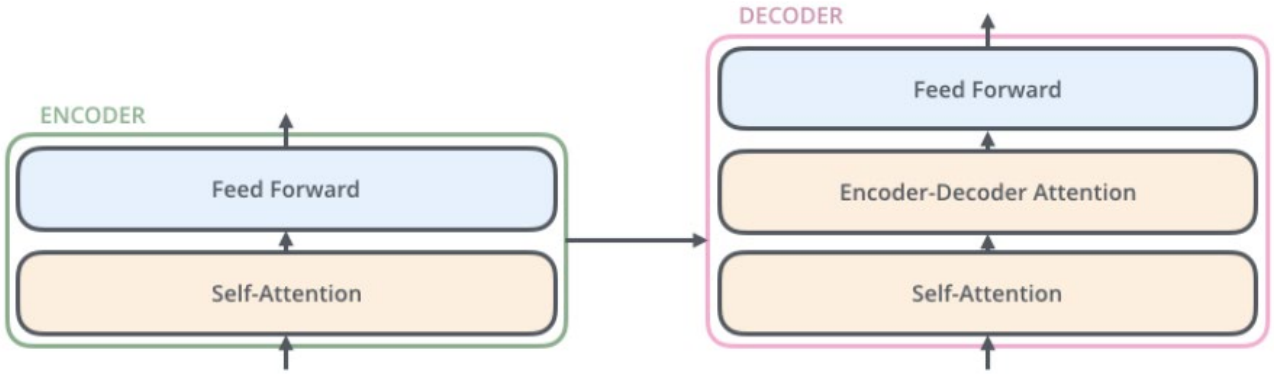
	IoU0.6 Acc	IoU0.9 Acc	Traj Cls Acc
Single	39.1%	6.5%	85.0%
FT transfer	18.5%	3.6%	74.3%
FT pretrained + finetune	43.5%	8.7%	89.8%
FHN transfer	25.2%	7%	78.8%
FHN pretrained + finetune	46.2%	9.2%	90.5%

Sample_len_avg: 68.7 frames
Label_len_avg: 15.5 frames

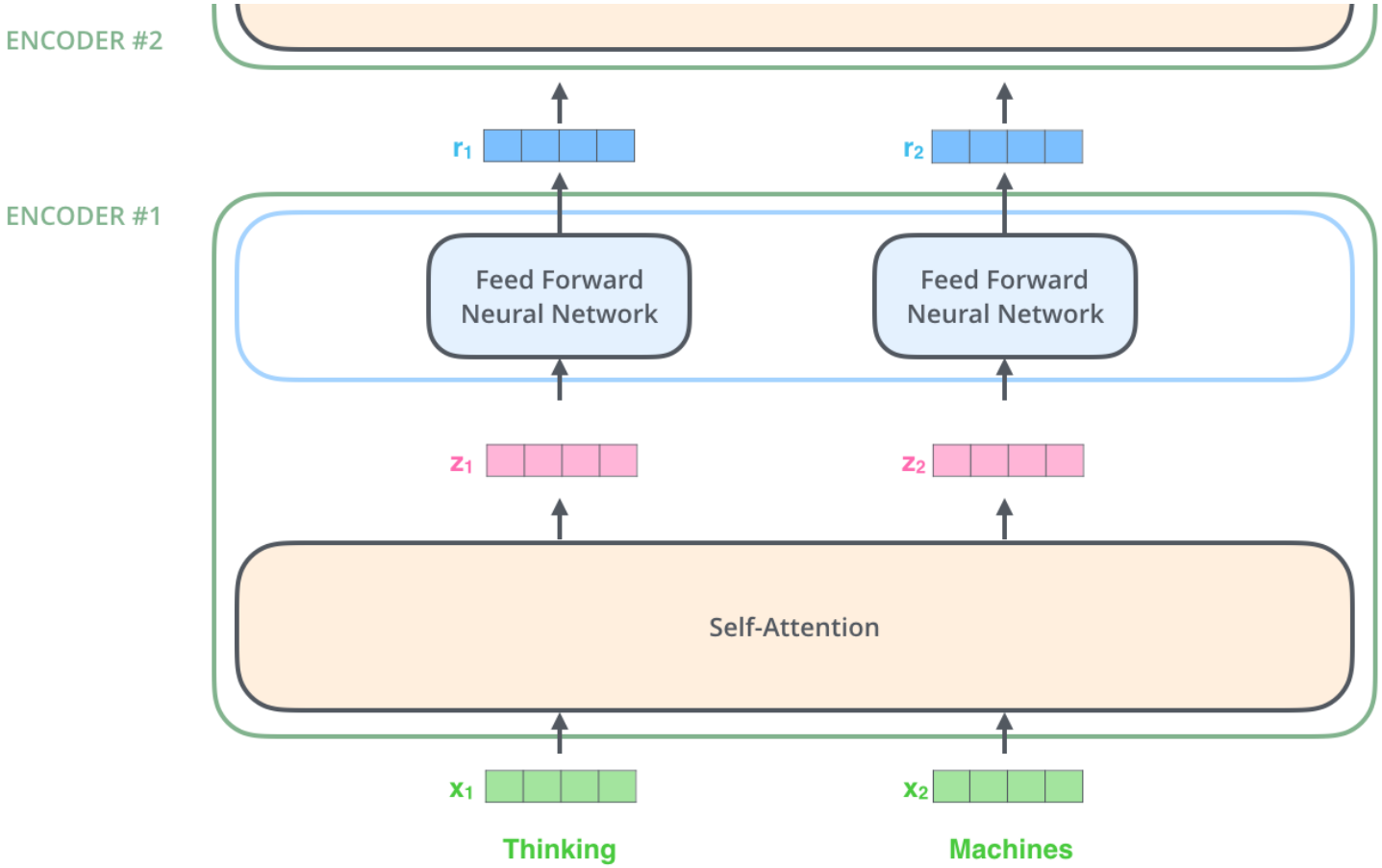


BERT (Overall Structure)

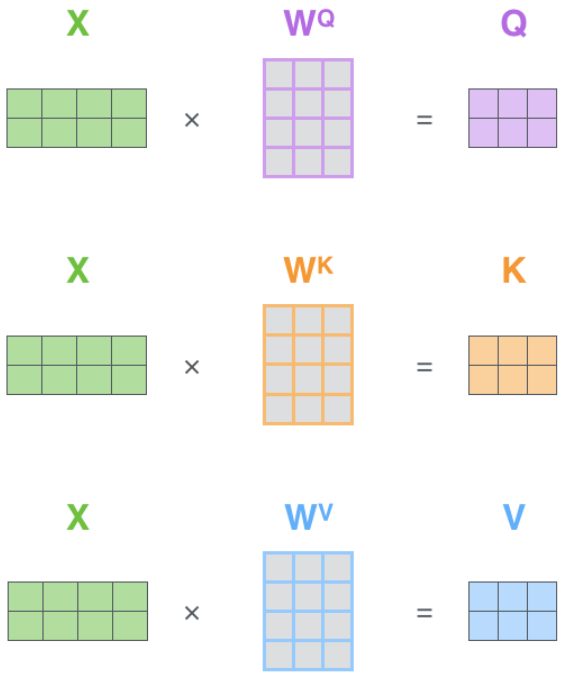


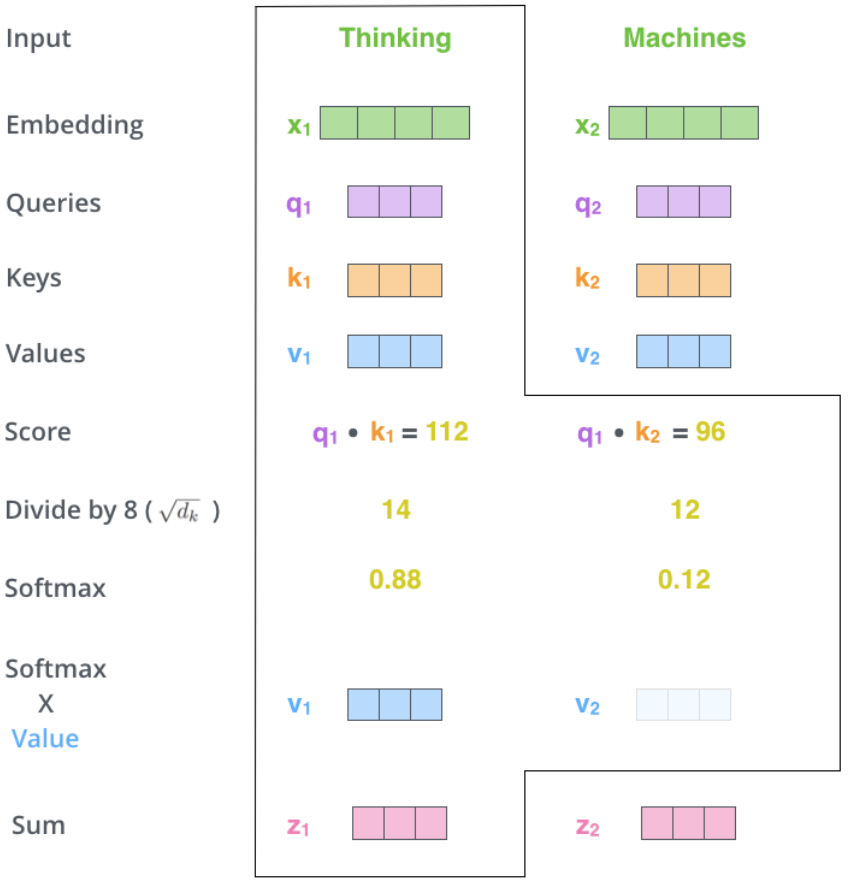


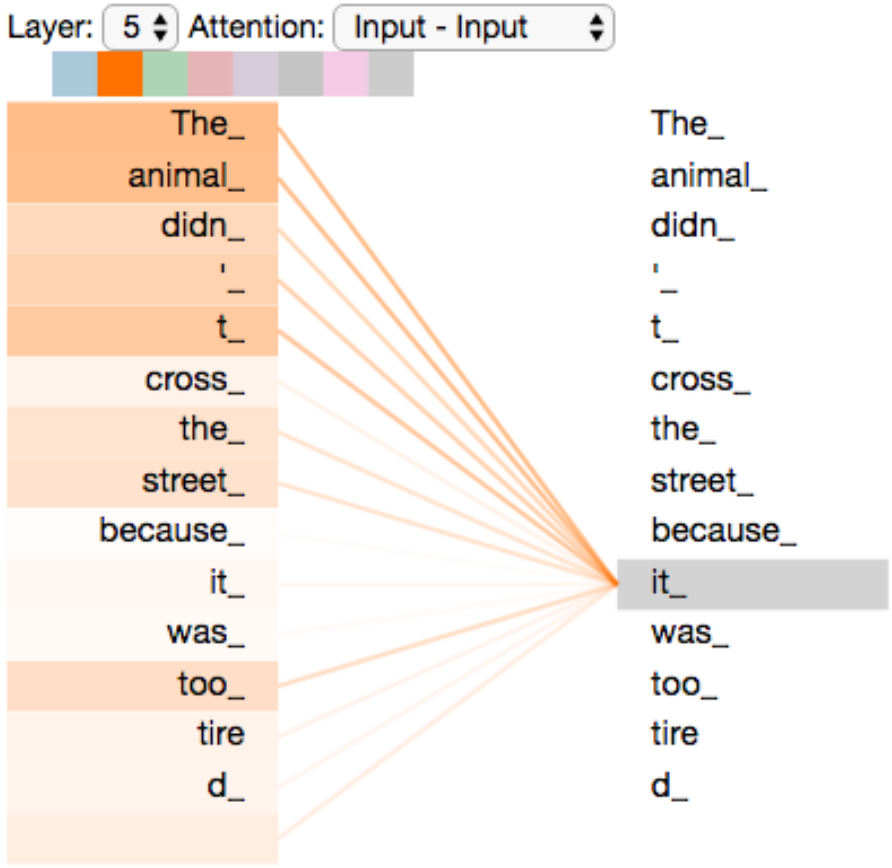
BERT (Decoder)



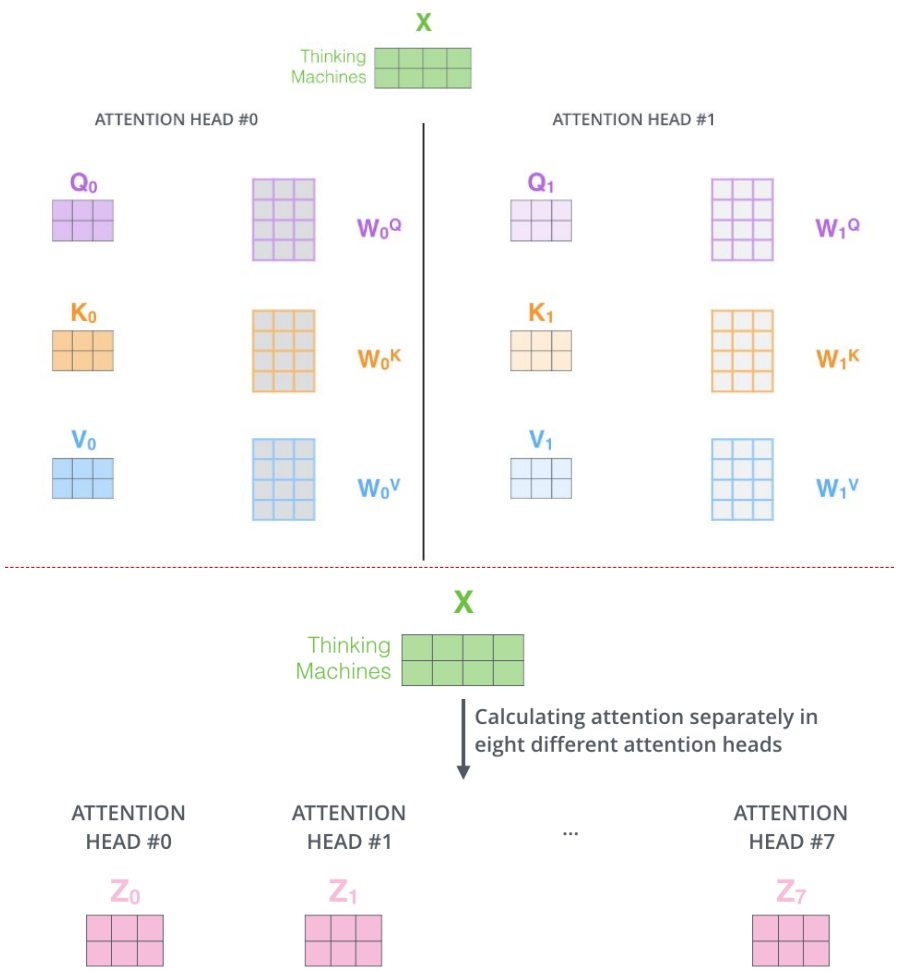
BERT (Self-attention)







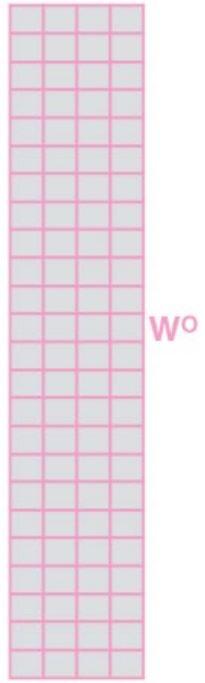
BERT (Multi-head)



1) Concatenate all the attention heads

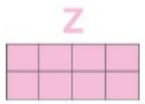
2) Multiply with a weight matrix W^O that was trained jointly with the model

X

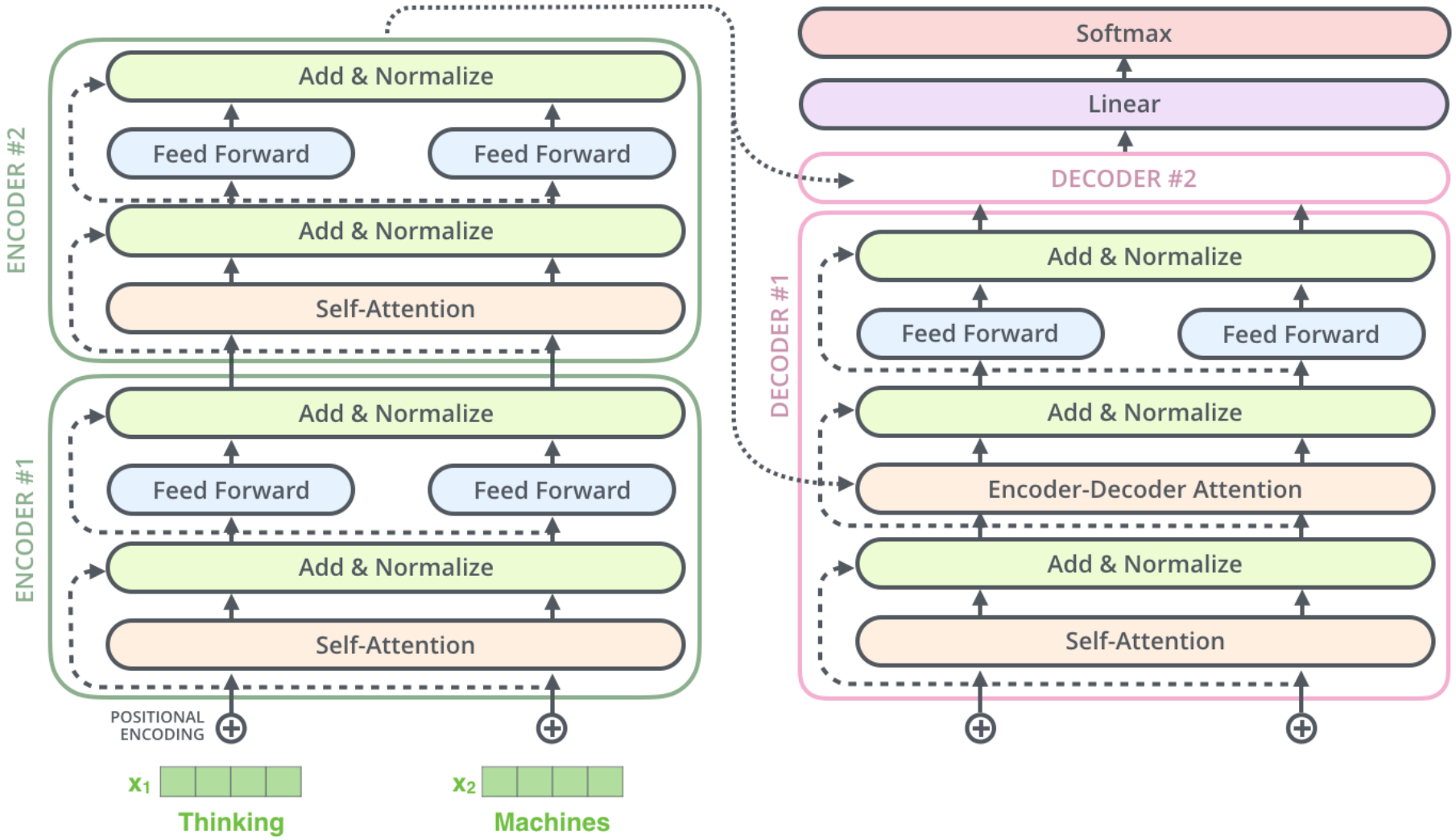


3) The result would be the Z matrix that captures information from all the attention heads. We can send this forward to the FFNN

=



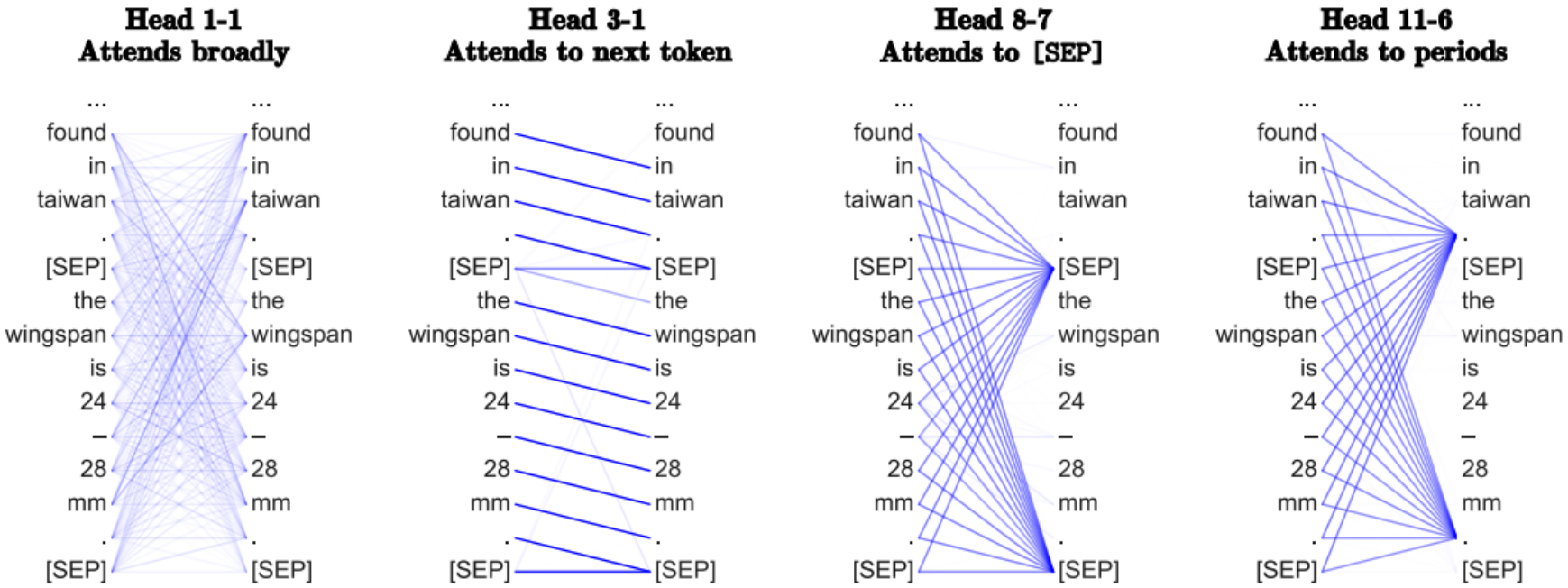
BERT (Decoder)

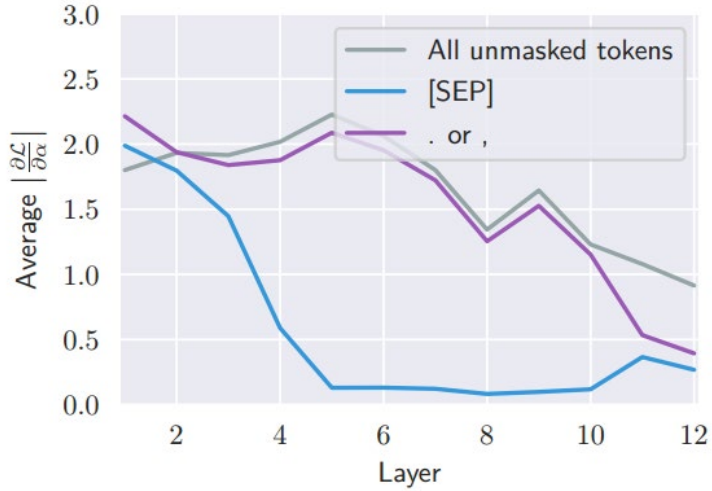


Some experiments

- UDA: too much hyperparameters -> hard to tune; train slowly; no improvements yet
- Encoder-Decoder: slower converging; 2% improvements on IoU0.6
- Relative Features with Encoder-Decoder (in training)
- Rotate entire trajectories: according to the system speed of the first step (in training)

BERT Visualization [Kevin Clark et. al. 2019]

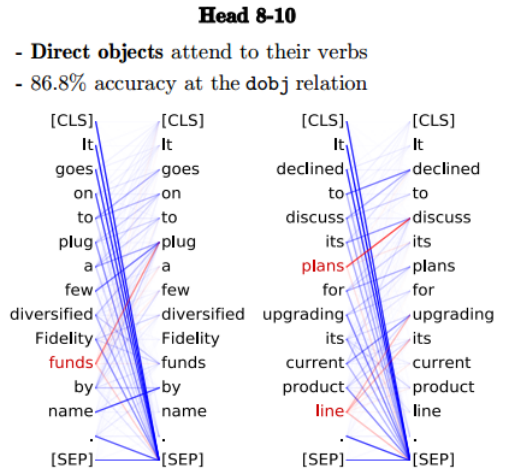


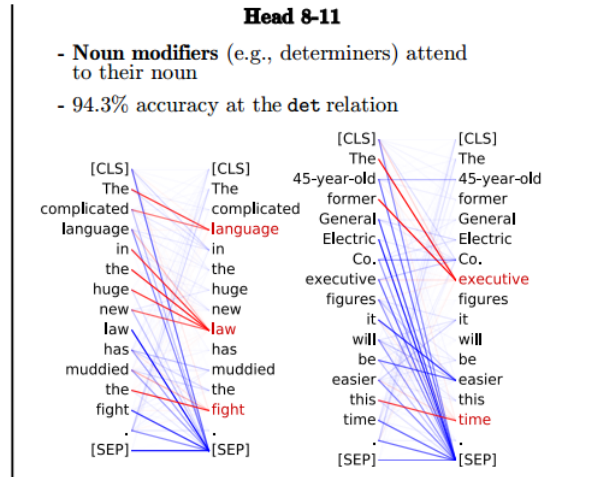


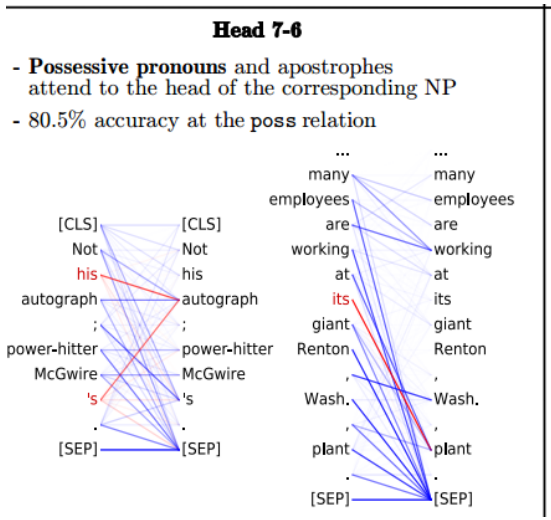
- Relative Position
- [Sep] -> Key? no-op!

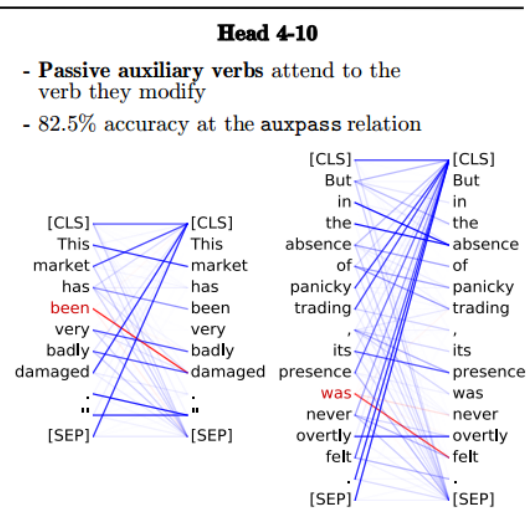
BERT Visualization [Kevin Clark et. al. 2019]

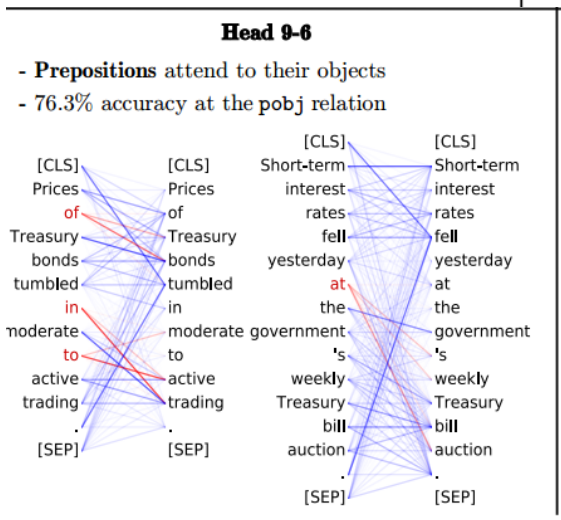
- Dependency Syntax
- No heads can do well at syntax over
- Certain attention heads specialize to specific dependency relations
- Even for coreferent

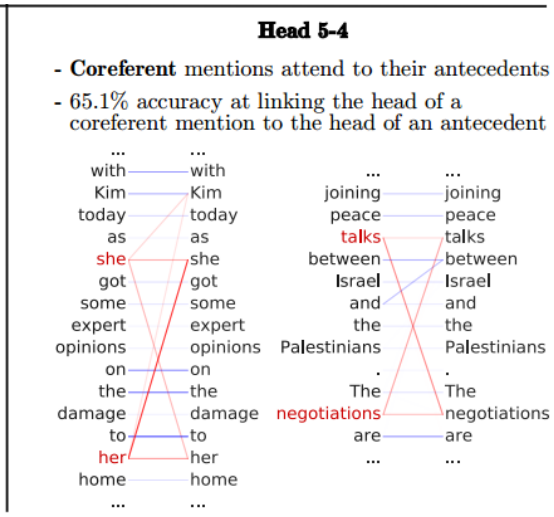












Next step

- Explanation (Single point -> Snippet?)

- More datasets