Vehicle Interaction Learning

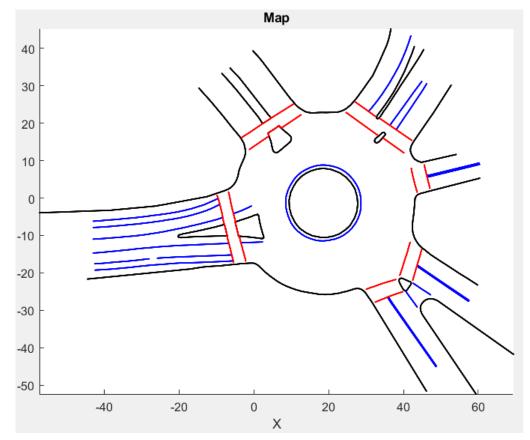
Xiaosong Jia

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

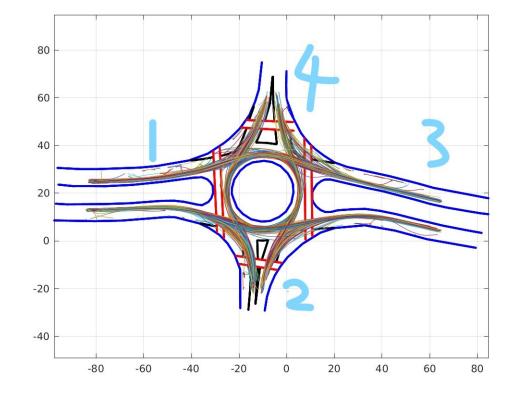
	IoU0.6 Acc	IoU0.9 Acc	Traj Cls Acc
HighD	94.8%	32.5%	99.1%
NGSIM	82.5%	18.9%	89.2%
FT	82.3%	35.1%	86.3%

USA_Roundabout_SR

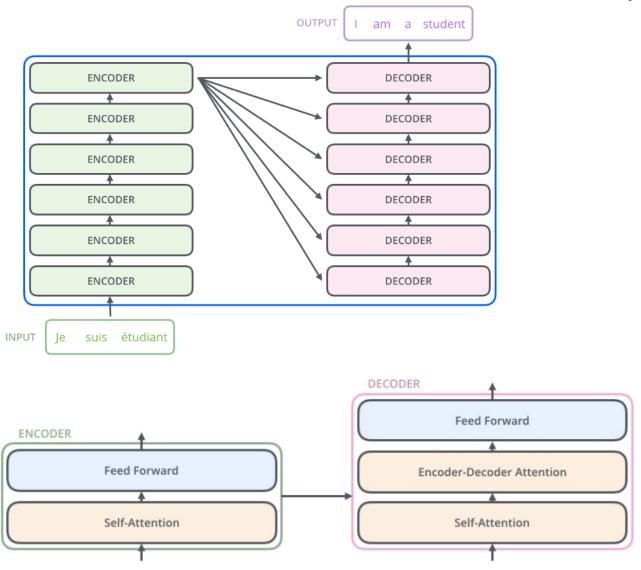
	Positive	Negative
Stop sign	22	2
TTC	213	124
Total	235	126

	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%

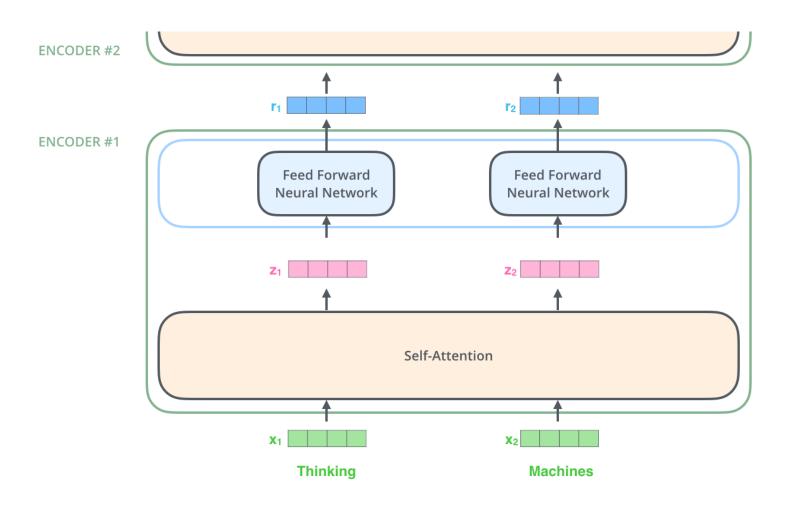
Sample_len_avg: 68.7 frames Label_len_avg: 15.5 frames



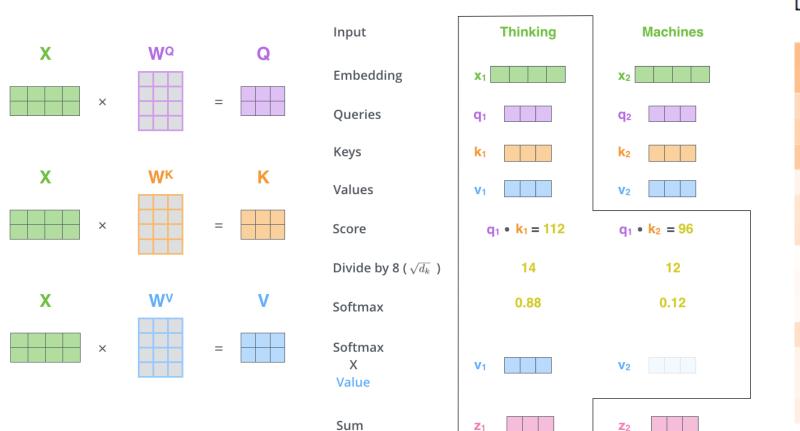
BERT (Overall Structure)

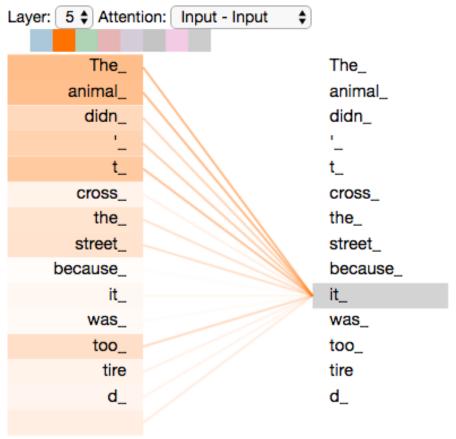


BERT (Decoder)

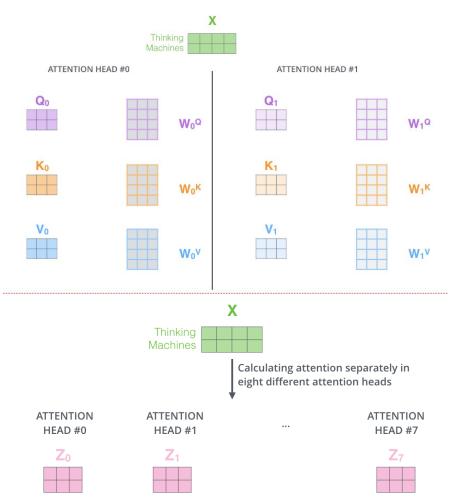


BERT (Self-attention)





BERT (Multi-head)







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

X

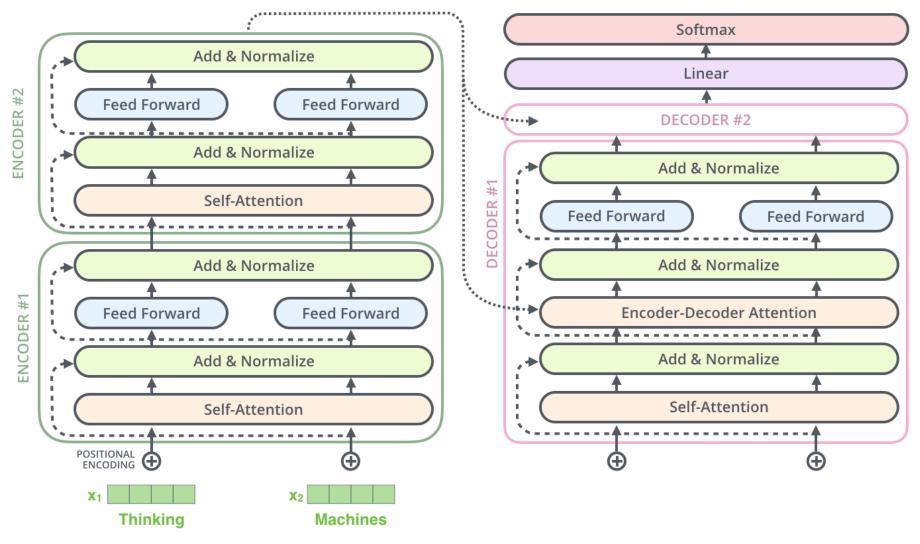
3) The result would be the $\mathbb Z$ matrix that captures information from all the attention heads. We can send this forward to the FFNN







BERT (Decoder)



Some experiments

 UDA: too much hyperparameters -> hard to tune; train slowly; no improvements yet

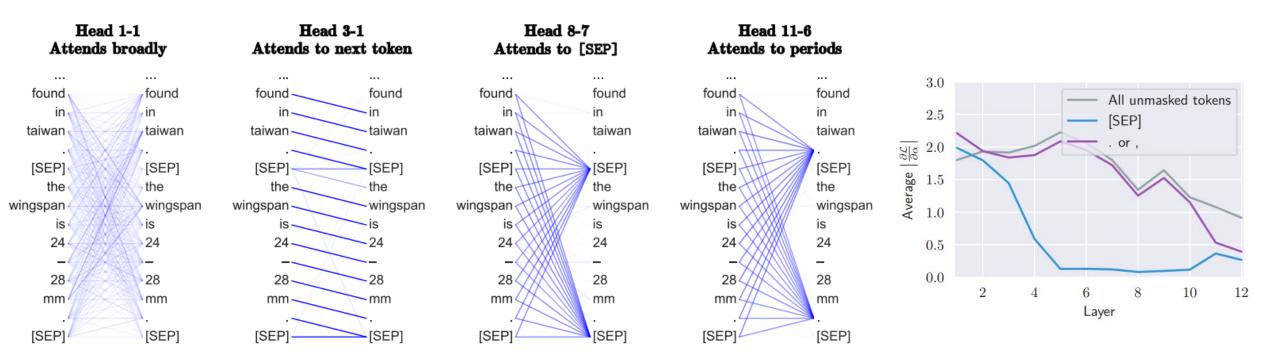
• Encoder-Decoder: slower converging; 2% improvements on IoU0.6

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!

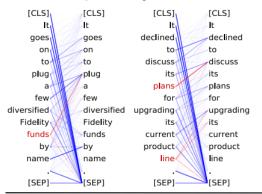


BERT Visualization [Kevin Clark et. al. 2019]

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

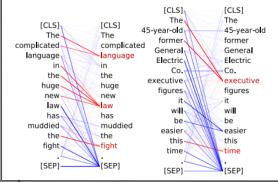
Head 8-10

- Direct objects attend to their verbs
- 86.8% accuracy at the dobj relation



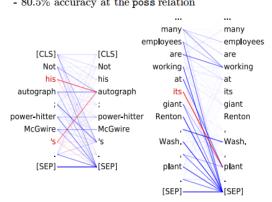
Head 8-11

- Noun modifiers (e.g., determiners) attend to their noun
- 94.3% accuracy at the det relation



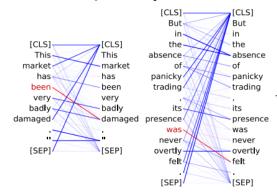
Head 7-6

- Possessive pronouns and apostrophes attend to the head of the corresponding NP
- 80.5% accuracy at the poss relation



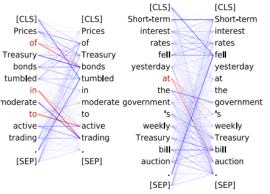
Head 4-10

- Passive auxiliary verbs attend to the verb they modify
- 82.5% accuracy at the auxpass relation



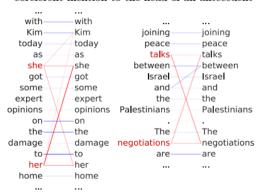
Head 9-6

- Prepositions attend to their objects
- 76.3% accuracy at the pobj relation



Head 5-4

- Coreferent mentions attend to their antecedents
- 65.1% accuracy at linking the head of a coreferent mention to the head of an antecedent



Next step

Explanation (Single point -> Snippet?)

More datasets

```
OUSA_FT ONGSIM ODEU_OF
OUSA_GL OUSA_MA OEUR_VA
OHighD OUSA_EPO OUSA_EP1
OUSA_EP OUSA_SR OUSA_CS
OCHN_LN OCHN_ZS
```