

Kuan Yu kuanyu@uni-potsdam.de

Maya Angelova maya.angelova@protonmail.com

Philipp Schoneville schoneville@uni-potsdam.de


Sonu Rauniyar rauniyar@uni-potsdam.de

June 27, 2018

¹*explicare*: to explain, to unfold; *char*: character; *arr*: array.

aligned sentences

- ▶ source: standard english wikipedia
- ▶ target: simple english wikipedia

²<http://ssli.ee.washington.edu/tial/projects/simplification/> 

character-level modelling

consider only the top 256 characters out of ~ 3000

- ▶ more robust
 - ▶ rare characters make up only 0.03% of the text
 - ▶ no special treatment for large numbers and named entities
 - ▶ may learn morphology
- ▶ less preprocessing
- ▶ easier applicable to other languages

neural network translator

S : source alphabet

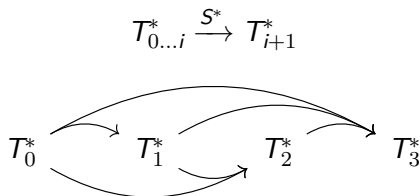
T : target alphabet

$$m : S^* \rightarrow T^*$$

$$= S^* \rightarrow T_{0\dots i}^* \rightarrow T_{i+1}^*$$

$$= T_{0\dots i}^* \xrightarrow{S^*} T_{i+1}^*$$

autoregressive network



growing number of arrows means

- ▶ growing number of parameters
 - ▶ parameter sharing (convolution, recurrent, attention)
- ▶ growing number of inputs
 - ▶ limit input field (convolution)
 - ▶ input aggregating (recurrent)
 - ▶ input averaging (attention)

transformer³

all attention

- ▶ no limited input field
- ▶ no information bottleneck
- ▶ no hidden to hidden connection
 - ▶ can be trained with teacher forcing
 - ▶ highly parallelizable

³<https://arxiv.org/abs/1706.03762>

current status

done

- ▶ data cleanup
- ▶ model implementation

todo

- ▶ introspection
- ▶ optimization

results

- ▶ ~200k training instances
- ▶ ~2k validation instances
- ▶ ~90% training accuracy (teacher forcing)
- ▶ ~29% blue score (autoregressive)

good results

source: In fact , they are different things .

target: These words mean things
that are a little different .

output: They are different things .

source: In more recent years ,
he has played a metal saxophone .

target: Now he plays a metal saxophone .

output: He has played a metal saxophone .

source: With one huge blow from his olive-wood club ,
Hercules killed the watchdog .

target: Herakles killed her .

output: Hercules killed the watchdog .

mystery

~88% of output sentences are exact copies of source sentences

- ▶ only when used autoregressively
- ▶ target and source don't share character inventory
- ▶ target and source don't share time steps

dual mystery

a deeper model completely ignores the source sentences

- ▶ unable to condition on the source
- ▶ becomes an autoencoder for the target (teacher forcing)
- ▶ always produces the same output (autoregressive)

future plan

- ▶ attention visualization
 - ▶ to solve the mystery
 - ▶ to understand what the model does
- ▶ autoregressive training
 - ▶ training with its own output
 - ▶ backprop through time
 - ▶ mean field approximation
- ▶ encoder pretraining
 - ▶ to solve the dual mystery
- ▶ decoding
 - ▶ beam search
 - ▶ soft predictions