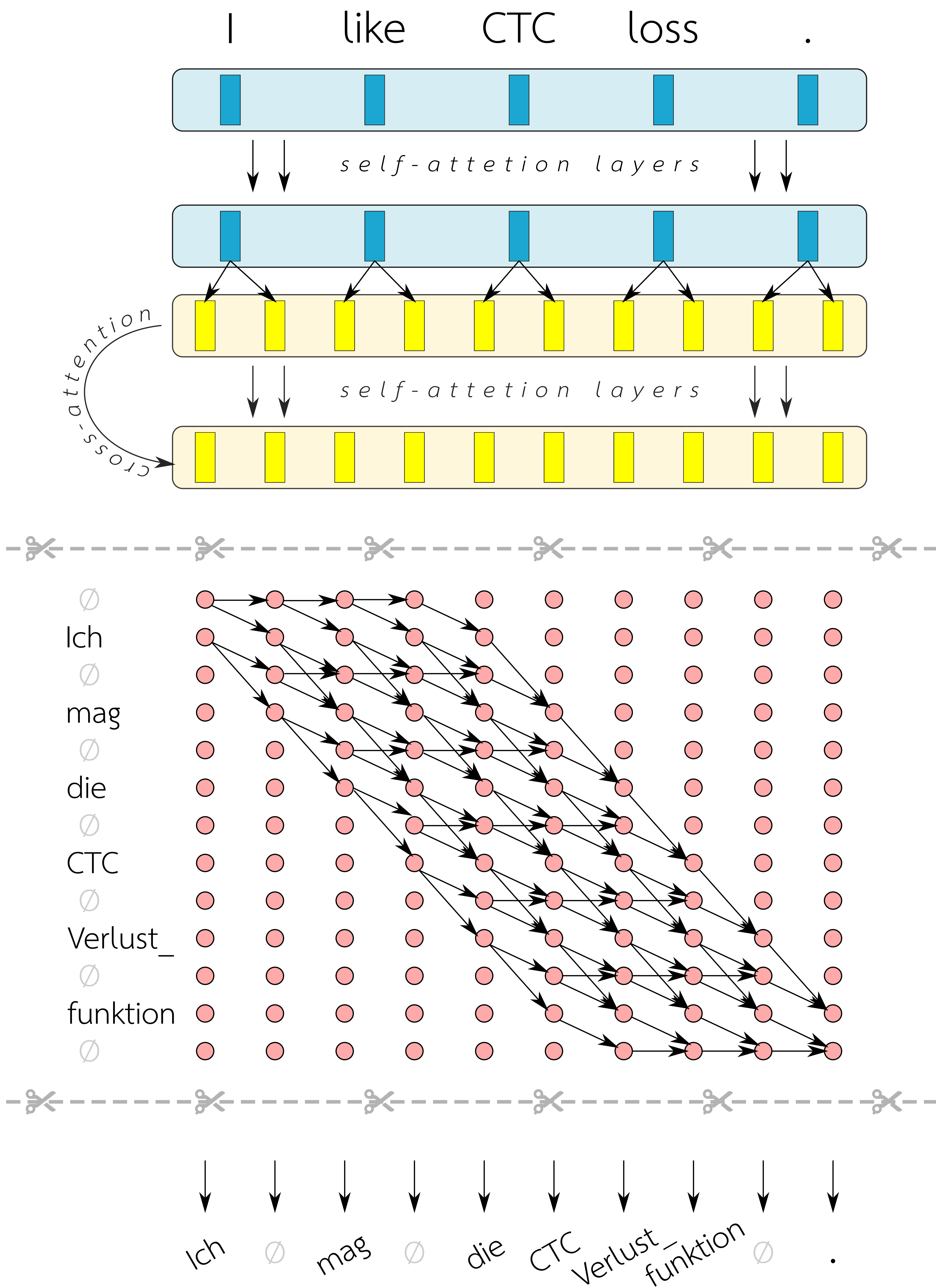


End-to-End Non-Autoregressive Neural Machine Translation with Connectionist Temporal Classification

J. Libovický, J. Helcl

Charles University



	WMT 16		WMT 14		WMT 15	
	en-ro	ro-en	en-de	de-en	en-de	de-en
autoregressive baseline, beam 1	21.19	29.64	22.94	28.58	25.12	28.89
Ge et al. (2017), greedy	27.29	29.06	17.69	21.47	-	-
Lee et al. (2018), 1 iteration	24.45	23.73	-	-	12.65	14.48
deep encoder	17.33	22.85	12.21	12.53	13.14	18.34
+ weight averaging	18.47	24.68	14.65	16.72	16.74	18.47
encoder-decoder	18.51	22.37	13.29	17.98	16.01	19.55
+ weight averaging	19.54	24.67	16.56	18.64	19.46	21.74
encoder-decoder with pos. encoding	18.13	22.75	12.51	11.35	15.35	19.30
+ weight averaging	19.31	24.21	17.37	18.07	20.30	19.64

