Dream team NLP speech

Tatiana Gaintseva et al (Aleksandr Mikhailov, Vyacheslav Krakhotin, Yuriy Gorishniy, Sergey Dzhunusov, Ivan Baskov)
25.10.2018, Moscow, Princeton

Google's Neural Machine Translation System

|	PBMT	GNMT	Human	Improvement
En->Es	4.885	5.428	5.504	87%
En->Fr	4.932	5.295	5.496	64%
En->Ch	4.035	4.594	4.987	58%
Es->En	4.872	5.187	5.372	63%
Fr->En	5.046	5.343	5.404	83%
Ch->En	3.694	4.263	4.636	60%

Human evaluation results:
- 6 - perfect translation
- 4 - almost perfect but with some grammar errors
- 2 - preserve some meaning but missing parts

Achieving Human Parity on Automatic Chinese to English News Translation

Translation	%	z
Model 6	69.0	0.237
Human translation from scratch	68.5	0.220
Model 5	68.9	0.216
Model 4	68.6	0.211
Google Translate, post-edited	67.3	0.141
Best model from WMT17	62.3	-0.094
Reference translation from WMT17	62.1	-0.115
MS Translator	56.0	-0.398
Google Translate	54.1	-0.468

Towards Robust Neural Machine Translation
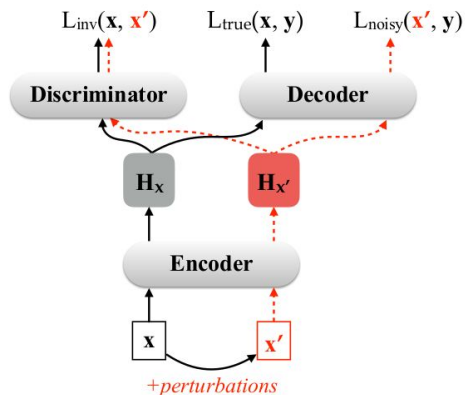
- **Swap**: We randomly choose N positions from a sentence and then swap the chosen words with their right neighbours.

- **Replacement**: We randomly replace sampled words in the sentence with other words.

- **Deletion**: We randomly delete N words from each sentence in the dataset.

Context-Aware Neural Machine Translation Learns Anaphora Resolution

model	BLEU
baseline	29.46
concatenation (previous sentence)	29.53
context encoder (previous sentence)	**30.14**
context encoder (next sentence)	29.31
context encoder (random context)	29.69

type	N	baseline	our model	diff.
masc.	2509	26.9	27.2	**+0.3**
fem.	2403	21.8	26.6	**+4.8**
neuter	862	22.1	24.0	**+1.9**
plural	1141	18.2	22.5	**+4.3**

Table 5: BLEU for test sets of pronoun "it" having a nominal antecedent in context sentence. N: number of examples in the test set.