Moses in a Neural MT World

Hieu Hoang

MT Marathon 2017

Dayton, Ohio

SMT ~~Moses~~ in a Neural MT World

Hieu Hoang

MT Marathon 2017

Dayton, Ohio

Why work on SMT?

- It’s still being used
- Competitive quality
  - Some language pairs
  - Low resource languages
  - Out-of-domain
- Translation speed
- Hybridization

It’s still used

- EU, WIPO, AutoDesk, Adobe, Amazon, eBay
- All use Moses

- Google
- 20 language pairs uses NMT
- 82 more to go!

- Bing
- 11 language pairs uses NMT
- 50 more to go!

Competitive Quality

United Nations Corpus

	ar-en	ar-es	ar-fr	ar-ru	ar-zh	en-ar	en-es	en-fr	en-ru	en-zh
pb	53.07	49.77	42.80	36.00	31.58	41.96	61.26	50.09	43.25	37.84
NMT	**55.97**	**52.16**	**44.67**	**38.77**	**39.73**	**45.07**	**62.12**	**51.46**	**45.27**	**46.82**

	es-ar	es-en	es-fr	es-ru	es-zh	fr-ar	fr-en	fr-es	fr-ru	fr-zh
pb	38.13	59.89	**49.76**	39.69	31.27	34.43	52.22	**52.44**	36.48	29.98
NMT	**39.69**	**61.13**	49.75	**40.96**	**41.21**	**35.53**	**52.38**	51.78	**37.26**	**37.37**

	ru-ar	ru-en	ru-es	ru-fr	ru-zh	zh-ar	zh-en	zh-es	zh-fr	zh-ru
pb	34.43	**52.59**	49.61	43.37	32.63	28.02	42.97	39.64	34.42	29.57
NMT	**35.73**	52.35	**50.45**	**44.14**	**39.42**	**35.38**	**51.77**	**48.29**	**42.00**	**36.41**

(Junczys-Dowmunt et al, 2016)

Translation Quality

en-de

	BLEU
PB	25.8
Hierarchical	24.6
NMT	**31.1**

(Bentivogli et al, 2016)

Translation Quality

en-de

	BLEU
uedin-nmt	34.2
metamind	32.3
NYU-UMontreal	30.8
cambridge	30.6
uedin-syntax	30.6
KIT/LIMSI	29.1
KIT	29.0
uedin-pbmt	28.4
jhu-syntax	26.6

de-en

	BLEU
uedin-nmt	38.6
uedin-pbmt	35.1
jhu-pbmt	34.5
uedin-syntax	34.4
KIT	33.9
jhu-syntax	31.0

- **NMT**
- **SMT**

(Sennrich, MTMA 2016)

Phrase-based v. Hierarchical SMT

en-jp

	BLEU
PB	29.8
Hierarchical	**32.56**
NMT (LSTM)	32.19

(Zhu, 2015)

zh-en & en-zh

	zh-en	en-zh
pb	43.0	37.84
hierarchical	47.3	41.7
NMT	**51.8**	**46.82**

(Junczys-Dowmunt et al, 2016)

Low-Resource Languages

	SMT (Syntax)	NMT
Hausa	**23.7**	16.8
Turkish	**20.4**	11.4
Uzbek	**17.9**	10.7
Urdu	**17.9**	5.2

(Zoph et al, 2016)

	SMT (PB)	NMT (GroundHog)
Mongolian-Chinese	**29.48**	27.32

(Wu et al, 2016)

Low-Resource Languages

- Loreili/Phi results

Out-of-Domain

Catalan-Spanish

	In-Domain	Out-of-Domain
Rule-based	75.20	50.53
Phrase-based	81.80	**57.20**
Neural	**83.01**	52.10

(Costa-jussa, 2017)

Speed

	Words / sec
Moses (16 cores)	455.3
Nematus (1 GPU)	268.4
AmuNMT (1 GPU)	**864.7**

(Junczys-Dowmunt et al, 2016)

Moses2

Speed / Quality trade-off

- Compare
  - Moses/Moses2 with different cube-pruning setting
  - Amun with different beams

Hybridization

Neural network in SMT

Re-rank n-best list

Schwenk

(Bulgarian/Dutch guy)

Feature function in SMT decoding

Devlin

(Oxford)