

HNMT – Helsinki Neural Machine Translation

Robert Östling and Jörg Tiedemann

2016-09-12

In brief

- Our research mostly concerns:

In brief

- Our research mostly concerns:
 - morphologically complex languages

In brief

- Our research mostly concerns:
 - morphologically complex languages
 - low-resource languages

In brief

- Our research mostly concerns:
 - morphologically complex languages
 - low-resource languages
- Two releases today:

In brief

- Our research mostly concerns:
 - morphologically complex languages
 - low-resource languages
- Two releases today:
 - HNMT – our experimental NMT system
<https://github.com/robertostling/hnmt>

In brief

- Our research mostly concerns:
 - morphologically complex languages
 - low-resource languages
- Two releases today:
 - HNMT – our experimental NMT system
<https://github.com/robertostling/hnmt>
 - BNAS – our experimental ANN library
<https://github.com/robertostling/bnas>

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM
 - character biLSTM

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM
 - character biLSTM
 - word biLSTM with character biLSTM for <UNK>

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM
 - character biLSTM
 - word biLSTM with character biLSTM for <UNK>
- Decoders (pick one):

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM
 - character biLSTM
 - word biLSTM with character biLSTM for <UNK>
- Decoders (pick one):
 - word LSTM

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM
 - character biLSTM
 - word biLSTM with character biLSTM for <UNK>
- Decoders (pick one):
 - word LSTM
 - character LSTM

The model(s)

<blink>**Work in progress**</blink>

- Morphology \implies character-based/hybrid approach
- We want to experiment with different architectures
- We mostly follow Luong & Manning (2016)
- Encoders (pick one):
 - word biLSTM
 - character biLSTM
 - word biLSTM with character biLSTM for <UNK>
- Decoders (pick one):
 - word LSTM
 - character LSTM
 - TODO: word LSTM with character LSTM for <UNK>

Low-resource languages

- NMT is data-hungry, how to translate low-resource languages?

Low-resource languages

- NMT is data-hungry, how to translate low-resource languages?
- Could Bayesian priors help?

Low-resource languages

- NMT is data-hungry, how to translate low-resource languages?
- Could Bayesian priors help?
- HNMT can use *efmaral* to provide auxiliary attention loss
<https://github.com/robertostling/efmaral>

Low-resource languages

- NMT is data-hungry, how to translate low-resource languages?
- Could Bayesian priors help?
- HNMT can use *efmaral* to provide auxiliary attention loss
<https://github.com/robertostling/efmaral>
- We don't know yet how well this works...

Low-resource languages

- NMT is data-hungry, how to translate low-resource languages?
- Could Bayesian priors help?
- HNMT can use *efmaral* to provide auxiliary attention loss
<https://github.com/robertostling/efmaral>
- We don't know yet how well this works...
- Question to the audience: other ideas?

Preliminary experiments (novel compounds in blue)

source The Commission's 2007 **enlargement strategy paper** should send a clear signal regarding the firm commitment toward the countries with which **accession negotiations** have started or the countries with **accession prospects**.

HNMT Kommissionens **utvidgningsstrategidokument** 2007 bör sända en tydlig signal om det fasta åtagandet mot de länder som **anslutningsförhandlingarna** har inlett eller länderna med **anslutningsframtidutsikter**.

reference Kommissionens[*sic*] bör i sitt **strategidokument för utvidgningen** 2007 ge en tydlig signal om det bindande åtagandet gentemot länder som har pågående **anslutningsförhandlingar** med EU eller länder med **utsikter till anslutning**.

Preliminary experiments

source	Jag skulle vilja tacka honom för hans kollegiala samarbete och önska honom lycka i hans nya ämbete.
HNMT	I would like to thank him for his collegiative cooperation and wish him looking to his new office.
reference	I would like to thank him for his helpful and considerate cooperation and to wish him all the best for his new post.