

Interactive-Predictive Translation based on Multiple Word-Segments

Miguel Domingo, Álvaro Peris, Francisco Casacuberta

midobal@prhlt.upv.es, lvapeab@prhlt.upv.es, fcn@prhlt.upv.es

Pattern Recognition and Human Language Technology Research Centre
Universitat Politècnica de València

EAMT 2016

Riga, May 30, 2016

Outline

1. Motivation
2. Segment-Based IMT
3. Experiments
4. Conclusions

Outline

1. Motivation
2. Segment-Based IMT
3. Experiments
4. Conclusions

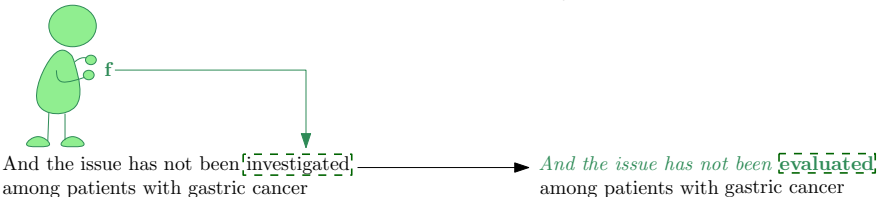
Motivation

- Machine Translation (MT) is still not good enough.
- Alternative to classical post-editing: Interactive-Predictive MT (IMT).
- Prefix-based IMT (Barrachina et al., 2009) was an interesting contribution to the field.
- Cumbersome phenomenon: the non-validated part of the sentence contains correct words.

Prefix-Based IMT

Source: Et la question n' a pas encore été évaluée chez les patients atteints de cancer gastrique

Target translation: And the issue has not been evaluated in gastric cancer patients



Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
----------	----------	--

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
	U	And the issue has not been evaluated among patients with gastric cancer

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been evaluated among patients with gastric cancer
	S	<i>And the issue has not been evaluated</i> with patients

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been evaluated among patients with gastric cancer
	S	<i>And the issue has not been evaluated</i> with patients
	U	<i>And the issue has not been evaluated</i> in patients

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been evaluated among patients with gastric cancer
	S	<i>And the issue has not been evaluated</i> with patients
2	U	<i>And the issue has not been evaluated</i> in patients
	S	<i>And the issue has not been evaluated in</i> patients

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been evaluated among patients with gastric cancer
	S	<i>And the issue has not been evaluated</i> with patients
2	U	<i>And the issue has not been evaluated</i> in patients
	S	<i>And the issue has not been evaluated in</i> patients
	U	<i>And the issue has not been evaluated in</i> gastric

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been evaluated among patients with gastric cancer
	S	<i>And the issue has not been evaluated</i> with patients
2	U	<i>And the issue has not been evaluated</i> in patients
	S	<i>And the issue has not been evaluated in</i> patients
3	U	<i>And the issue has not been evaluated in</i> gastric
	S	<i>And the issue has not been evaluated in gastric</i> cancer patients

Prefix-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been evaluated among patients with gastric cancer
	S	<i>And the issue has not been evaluated</i> with patients
2	U	<i>And the issue has not been evaluated</i> in patients
	S	<i>And the issue has not been evaluated in</i> patients
3	U	<i>And the issue has not been evaluated in</i> gastric
	S	<i>And the issue has not been evaluated in gastric</i> cancer patients
E	U	And the issue has not been evaluated in gastric cancer patients

Outline

1. Motivation
2. Segment-Based IMT
3. Experiments
4. Conclusions

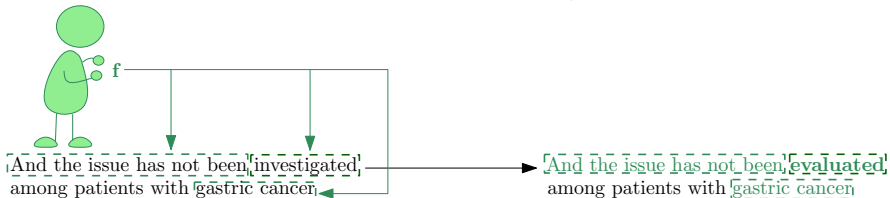
Segment-Based IMT

Goal: to develop a new IMT protocol which offers more freedom to the user, breaking down the prefix constraint.

- Select, remove, or replace segments of a translation hypothesis.
- New compatible hypothesis.

Source: Et la question n' a pas encore été évaluée chez les patients atteints de cancer gastrique

Target translation: And the issue has not been evaluated in gastric cancer patients



Segment-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

Segment-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
----------	----------	--

Segment-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
	U	And the issue has not been investigated among patients with gastric cancer

Segment-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
	U	<div style="border: 1px solid black; display: inline-block; padding: 2px;">And the issue has not been</div> investigated among patients with <div style="border: 1px solid black; display: inline-block; padding: 2px;">gastric cancer</div>
		<div style="border: 1px solid black; display: inline-block; padding: 2px;">And the issue has not been</div> evaluated among patients with <div style="border: 1px solid black; display: inline-block; padding: 2px;">gastric cancer</div>

Segment-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been investigated among patients with gastric cancer
	U	And the issue has not been evaluated among patients with gastric cancer
	S	And the issue has not been evaluated in gastric cancer patients

Segment-Based IMT Session

source (x): Et la question n' a pas encore été évaluée chez
les patients atteints de cancer gastrique

target translation (\hat{y}): And the issue has not been evaluated in gastric cancer patients

0	S	And the issue has not been investigated among patients with gastric cancer
1	U	And the issue has not been investigated among patients with gastric cancer
	S	And the issue has not been evaluated among patients with gastric cancer
	S	And the issue has not been evaluated in gastric cancer patients
E	U	And the issue has not been evaluated in gastric cancer patients

Statistical Framework

SMT (Brown et al., 1993):

$$\tilde{\mathbf{y}} = \arg \max_{\mathbf{y}} Pr(\mathbf{y} | \mathbf{x})$$

\mathbf{x} : source sentence.

\mathbf{y} : target sentence.

Prefix-based IMT (Barrachina et al., 2009):

$$\tilde{\mathbf{s}} = \arg \max_{\mathbf{s}} Pr(\mathbf{s} | \mathbf{x}, \hat{\mathbf{p}})$$

$\tilde{\mathbf{s}}$: suffix generated by the system.

$\hat{\mathbf{p}}$: validated prefix.

$$\tilde{\mathbf{y}} = \hat{\mathbf{p}}\tilde{\mathbf{s}}$$

Applying Bayes' rule:

$$\tilde{\mathbf{s}} = \arg \max_{\mathbf{s}} Pr(\hat{\mathbf{p}}, \mathbf{s} | \mathbf{x})$$

Search in the space of the translations, constrained by the prefix $\hat{\mathbf{p}}$.

Statistical Framework

Segment-based IMT:

$$\tilde{\mathbf{h}}_1, \dots, \tilde{\mathbf{h}}_N = \arg \max_{\mathbf{h}_1, \dots, \mathbf{h}_N} Pr(\hat{\mathbf{f}}_1 \mathbf{h}_1, \dots, \hat{\mathbf{f}}_N \mathbf{h}_N \mid \mathbf{x})$$

$\hat{\mathbf{f}}_1, \dots, \hat{\mathbf{f}}_N$: sequence of N segments validated by the user (feedback signal).

$\tilde{\mathbf{h}}_1, \dots, \tilde{\mathbf{h}}_N$: sequence of new translation segments.

$$\tilde{\mathbf{y}} = \hat{\mathbf{f}}_1 \tilde{\mathbf{h}}_1, \dots, \hat{\mathbf{f}}_N \tilde{\mathbf{h}}_N$$

Search in the space of the translation, constrained by the sequence of segments $\hat{\mathbf{f}}_1, \dots, \hat{\mathbf{f}}_N$.

Outline

1. Motivation
2. Segment-Based IMT
3. Experiments
4. Conclusions

Implementation

- Based on the Moses (Koehn et al., 2007) XML markup scheme.
- We force the translation of parts of a sentence into the decoder without changing the models.

And the issue has not been **evaluated** among patients with gastric cancer

```

<x translation = ‘‘And the issue has not been’’> Et la
question n ’ a pas encore été</x><wall/> <x translation
= ‘‘evaluated’’> évaluée</x><wall/>chez les patients
atteints de <x translation = ‘‘gastric cancer’’> cancer
gastrique</x><wall/>
    
```

Example of a sentence in XML

Implementation

An XML tag is created for each:

- **Validated segment:** the target words of the validated segment are linked to the corresponding source words*.
- **Deleted word:** a white translation is linked to the corresponding source words*.
- **Corrected word:** the new word is link to the corresponding source words*.

MT systems were trained with Moses standard configuration.

Prefix-based IMT systems were implemented using word graphs as described by Barrachina et al. (2009).

*Corresponding source words are obtained with an alignment model.

Corpora

EMEA (Tiedemann, 2009), **EU** (Barrachina et al., 2009), **TED** (Federico et al., 2011), and **Xerox** (Barrachina et al., 2009).

		EMEA (Fr/En)	EU (Es/En)	TED (Zh/En)	Xerox (Es/En)
Train	sentences	1.1M	214K	106.9K	55.6K
	tokens	14.3M/17.0M	6M/5.4M	1.9M/2.1M	750K/665K
	vocabulary	71K/80K	84K/70K	55K/41.7K	16.8K/14K
Dev.	sentences	500	400	934	1012
	tokens	12K/10K	12K/10K	21.5K/20.1K	16K/14.4K
	vocabulary	2.9K/2.7K	3K/2.7K	3.8K/3.2K	1.8K/1.6K
Test	sentences	1K	800	1.6K	1.1K
	tokens	27K/21K	23K/20K	33.2K/31.9K	10.1K/8.4K
	vocabulary	4.5K/4.5K	4.7K/4.2K	4.5K/3.7K	2K/1.9K

Metrics

Post-editing effort:

- **Word Stroke Ratio (WSR)** (Tomás and Casacuberta, 2006).
 - ▶ Correcting a word. (1 stroke.)
- **Mouse Action Ratio (MAR)** (Barrachina et al., 2009).
 - ▶ Correcting a word. (1 action.)
 - ▶ Validating a sentence. (1 action.)
 - ▶ Validating a segment. (2* actions.)
 - ▶ Deleting words between segments. (2* actions.)

Translation quality:

- **BiLingual Evaluation Understudy (BLEU)** (Papineni et al., 2002)

*1 action is enough for single word segments.

Evaluation

- Evaluation was carried out on a simulated environment.
- Corpora references were used as the user's desired translations.
- We assumed that the user always corrected the left-most wrong word.

Results

Prefix-based IMT in comparison with our segment-based proposal.

Corpus	Language	BLEU	Prefix-Based		Segment-Based	
			WSR (%)	MAR (%)	WSR (%)	MAR (%)
EMEA	Fr-En	31.3	57.8	12.4	34.4	18.8
	En-Fr	30.2	58.4	12.5	40.4	16.3
EU	Es-En	48.2	45.6	10.2	28.3	15.0
	En-Es	48.7	44.6	9.7	29.8	13.5
TED	Zh-En	11.7	83.1	22.4	54.1	28.3
	En-Zh	8.7	86.3	55.7	59.2	72.4
Xerox	Es-En	54.5	35.8	10.5	23.2	16.9
	En-Es	62.2	28.3	7.9	22.1	12.5

Results

- Substantial reduction of the typing effort (up to 29 points of WSR).
- Slight increase in the number of mouse actions (from 4 up to 6.5 points of MAR).

Outline

1. Motivation
2. Segment-Based IMT
3. Experiments
4. Conclusions

Conclusions

- New IMT approach that breaks down the prefix constraint.
- The user can select all correct words from each translation hypothesis.
- Post-editing effort effectively reduced in a simulated environment.
- Future work: experiments with real users.

References

- Barrachina, Sergio, Oliver Bender, Francisco Casacuberta, Jorge Civera, Elsa Cubel, Shahram Khadivi, Antonio Lagarda, Hermann Ney, Jesús Tomás, Enrique Vidal, and Juan-Miguel Vilar (2009). **Statistical Approaches to Computer-Assisted Translation**. In: *Computational Linguistics 35*, pp. 3–28.
- Brown, Peter F., Vincent J. Della Pietra, Stephen A. Della Pietra, and Robert L. Mercer (1993). **The Mathematics of Statistical Machine Translation: Parameter Estimation**. In: *Comput. Linguist.* 19.2, pp. 263–311.
- Federico, Marcello, Luisa Bentivogli, Michael Paul, and Sebastian Stüker (2011). **Overview of the IWSLT 2011 evaluation campaign**. In: *International Workshop on Spoken Language Translation*, pp. 11–27.
- Koehn, Philipp, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin, and Evan Herbst (2007). **Moses: Open Source Toolkit for Statistical Machine Translation**. In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 177–180.
- Tiedemann, Jörg (2009). **News from OPUS - A Collection of Multilingual Parallel Corpora with Tools and Interfaces**. In: *Recent Advances in Natural Language Processing. Vol. V*, pp. 237–248.
- Tomás, Jesús and Francisco Casacuberta (2006). **Statistical Phrase-Based Models for Interactive Computer-Assisted Translation**. In: *Proceedings of the Coling/ACL*, pp. 835–841.