

Machine Translation Slides

CS 4650/7650, Prof Jacob Eisenstein

April 2018

Translations, easy and hard

Sometimes, it's easy.

- ▶ Asha come pescado.
- ▶ Asha eats fish.

Translations, easy and hard

Sometimes, it's easy.

- ▶ Asha come pescado.
- ▶ Asha eats fish.

Word orders vary.

- ▶ Asha come pescado **crudo**.
- ▶ Asha eats **raw** fish.

Translations, easy and hard

Sometimes, it's easy.

- ▶ Asha come pescado.
- ▶ Asha eats fish.

Word orders vary.

- ▶ Asha come pescado **crudo**.
- ▶ Asha eats **raw** fish.

Information goes in different places.

- ▶ Asha no comió pescado.
- ▶ Asha didn't eat fish.

More translations

Language express different types of information.

- ▶ Asha eats **her** fish
- ▶ Asha mange **son** poisson

More translations

Language express different types of information.

- ▶ Asha eats **her** fish
- ▶ Asha mange **son** poisson

- ▶ Asha eats **her** apple
- ▶ Asha mange **sa** pomme

More translations

Language express different types of information.

- ▶ Asha eats **her** fish
- ▶ Asha mange **son** poisson

- ▶ Asha eats **her** apple
- ▶ Asha mange **sa** pomme

Different ways of saying the same thing.

- ▶ Asha likes fish.
- ▶ A Asha le gusta el pescado.

BLEU score

System	translation	p_1	p_2	p_3	p_4	BP	BLEU
<i>Reference</i>	<i>Vinay likes programming in Python</i>						
<i>Sys1</i>	<i>To Vinay it like to program Python</i>	$\frac{2}{7}$	0	0	0	1	.21
<i>Sys2</i>	<i>Vinay likes Python</i>	$\frac{3}{3}$	$\frac{1}{2}$	0	0	.51	.33
<i>Sys3</i>	<i>Vinay likes programming in his pajamas</i>	$\frac{4}{6}$	$\frac{3}{5}$	$\frac{2}{4}$	$\frac{1}{3}$	1	.76

Figure 1: BLEU scores

Performance by sentence length

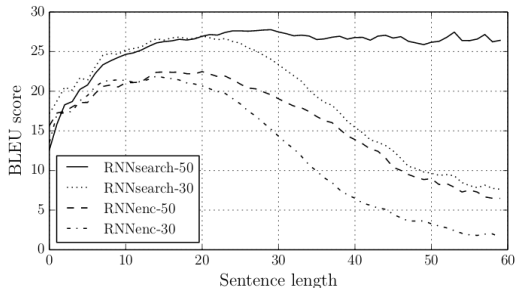


Figure 2: The BLEU scores of the generated translations on the test set with respect to the lengths of the sentences. The results are on the full test set which includes sentences having unknown words to the models.

Figure 2: Performance by sentence length, with and without attention

Attention in En-Fr translation

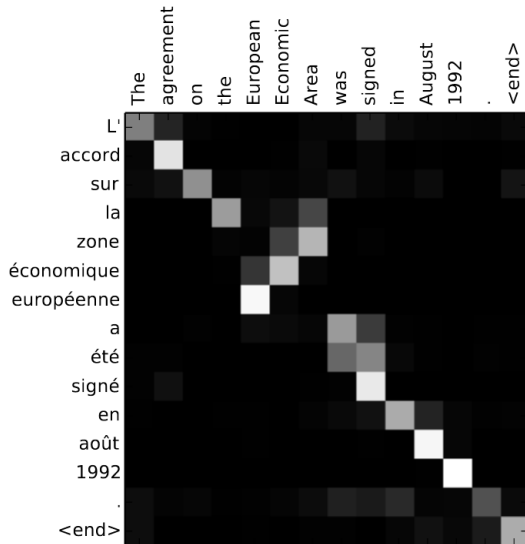


Figure 3: Example 1

Attention in En-Fr translation

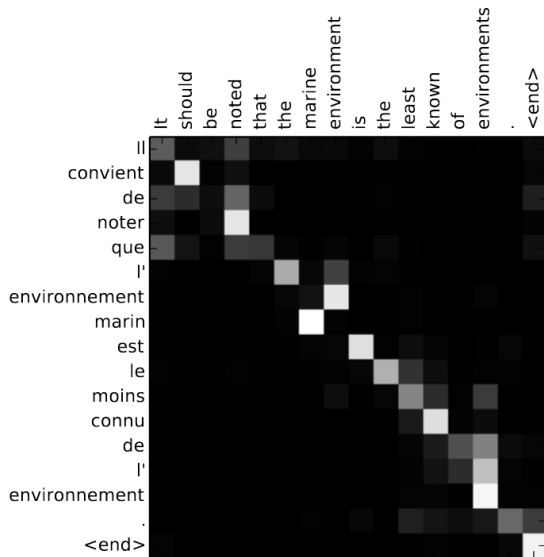


Figure 4: Example 2

OOV words

en: The ecotax portico in Pont-de-Buis, ... [truncated] ..., was taken down on Thursday morning

fr: Le portique écotaxe de Pont-de-Buis, ... [truncated] ..., a été démonté jeudi matin

nn: Le unk de unk à unk, ... [truncated] ..., a été pris le jeudi matin

Figure 1: **Example of the rare word problem** – An English source sentence (*en*), a human translation to French (*fr*), and a translation produced by one of our neural network systems (*nn*) before handling OOV words. We highlight words that are unknown to our model. The token unk indicates an OOV word. We also show a few important alignments between the pair of sentences.

Figure 5: Out-of-vocabulary words