Do Context-Aware Translation Models Pay the Right Attention?

Kayo Yin[1], Patrick Fernandes[1,2], Danish Pruthi[1], Aditi Chaudhary[1]
André F.T. Martins[2,3], Graham Neubig[1]

1 Carnegie Mellon University Language Technologies Institute

2 TÉCNICO LISBOA

3 Unbabel

Why is Context Important for Translation?

We'll have to get rid of that mole.

Why is Context Important for Translation?

Things could start to get dangerous if the ministers find out.
We'll have to get rid of that mole.

Why is Context Important for Translation?

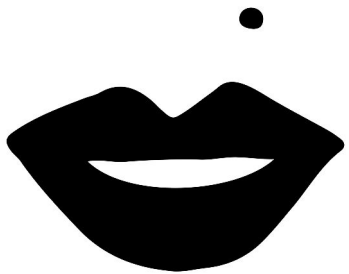
Things could start to get dangerous if the ministers find out.
We'll have to get rid of that mole.

Why is Context Important for Translation?

Could it be anything serious, Doctor?
We'll have to get rid of that mole.

Why is Context Important for Translation?

Could it be anything serious, Doctor?
We'll have to get rid of that mole.



Why is Context Important for Translation?

English:
Things could start to get dangerous if the ministers find out.
We'll have to get rid of that mole.

French:
Les choses pourraient commencer à devenir dangereuses si les ministres le découvraient.
Nous devrons nous débarrasser de cette taupe.

Why is Context Important for Translation?

English:
Could it be anything serious, Doctor?
We'll have to get rid of that mole.

French:
Serait-ce quelque chose de grave, docteur ?
Nous devrons nous débarrasser de ~~cette taupe~~.
cet grain de beauté

Why is Context Important for Translation?

English:
So you see how bad the implications are.
Yes, they are quite devastating.

French:
Vous voyez donc à quel point les implications sont mauvaises.
Oui, ils sont assez dévastateurs.

Why is Context Important for Translation?

English:
So you see how bad the implications are.
Yes, they are quite devastating.

French:
Vous voyez donc à quel point les implications sont mauvaises.
Oui, ~~ils~~ sont assez ~~dévastateurs.~~

elles dévastatrices

Context-Aware NMT

Source input	Have we got her report? Yes, it's in the infirmary already.
Context-aware NMT output	On dispose de son rapport? Oui, elle est déjà à l'infirmerie.

Context-Aware NMT

Source input	Have we got her report? Yes, it's in the infirmary already.
Context-aware NMT output	On dispose de son rapport? Oui, elle est déjà à l'infirmerie.

Context-Aware NMT

Source input	Have we got her report? Yes, it's in the infirmary already.
Context-aware NMT output	On dispose de son rapport? Oui, elle est déjà à l'infirmerie.

Context-Aware NMT

Source input

~~Have we got her report~~?
Yes, it's in the infirmary already.

Context-aware NMT output

~~On dispose de son rapport~~?
Oui, elle est déjà à l'infirmerie.

Outline

1. What context is useful during ambiguous translations?

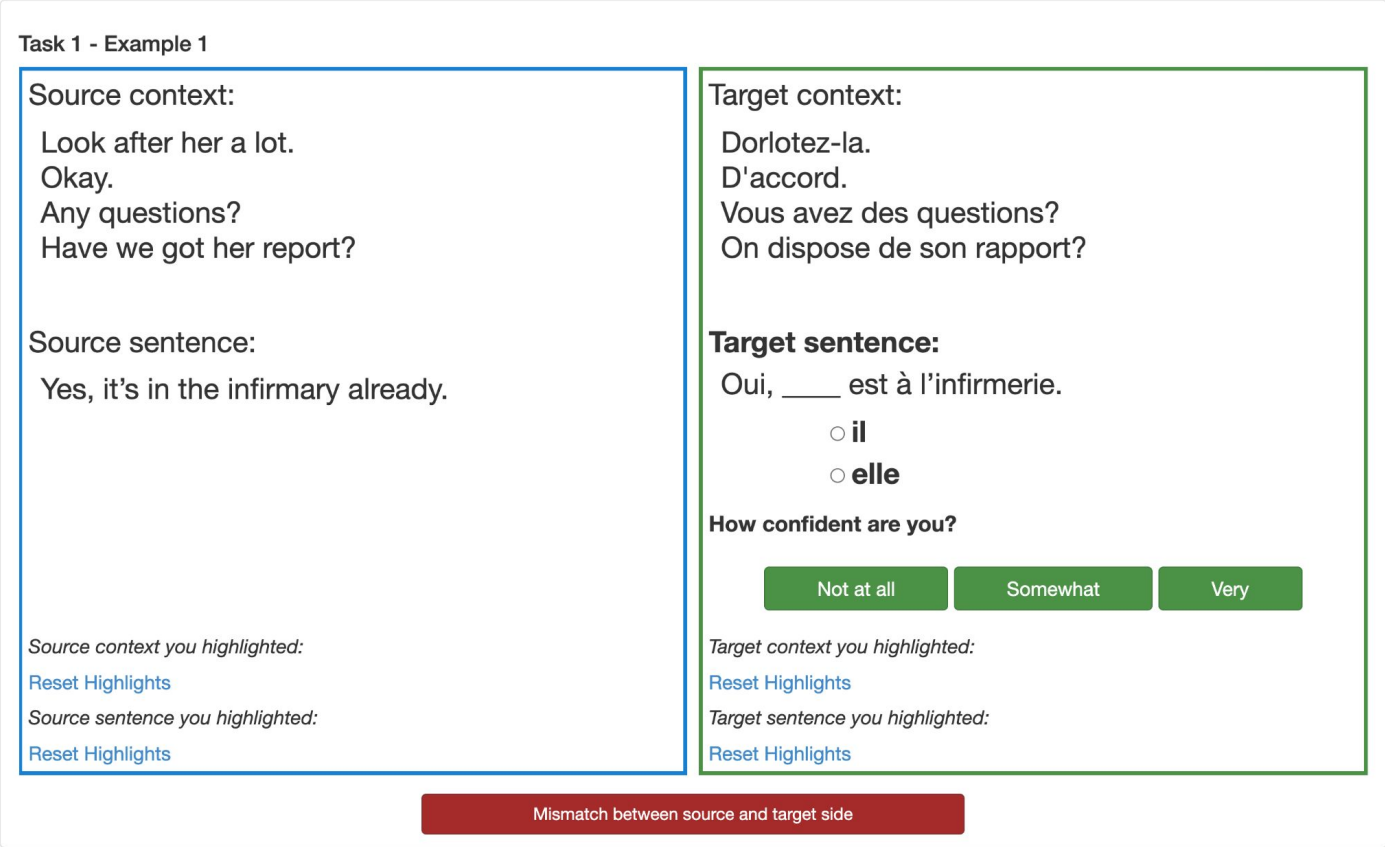
2. Are models paying attention to this context or not?

3. If not, can we encourage them to do so?

Outline

1. What context is useful during translation?
2. Are models paying attention to this context or not?
3. If not, can we encourage them to do so?

User Study


Task 1 - Example 1
Source context:
Look after her a lot.
Okay.
Any questions?
Have we got her report?
Source sentence:
Yes, it's in the infirmary already.
Source context you highlighted:
Reset Highlights
Source sentence you highlighted:
Reset Highlights
Target context:
Dorlotez-la.
D'accord.
Vous avez des questions?
On dispose de son rapport?
Target sentence:
Oui, ____ est à l'infirmerie.
il
elle
How confident are you?
Not at all
Somewhat
Very
Target context you highlighted:
Reset Highlights
Target sentence you highlighted:
Reset Highlights
Mismatch between source and target side

User Study

Task 1 - Example 1

Source context:

Look after her a lot.
Okay.
Any questions?
Have we got her report?

Source sentence:

Yes, it's in the infirmary already.

Source context you highlighted:

Reset Highlights

Source sentence you highlighted:

Reset Highlights

Target context:

Dorlotez-la.
D'accord.
Vous avez des questions?
On dispose de son rapport?

Target sentence:

Oui, ____ est à l'infirmerie.

- (•) **il**
- ( ) **elle**

How confident are you?

Not at all | Somewhat | Very

Target context you highlighted:

Reset Highlights

Target sentence you highlighted:

Reset Highlights

Mismatch between source and target side

User Study

Task 1 - Example 1

Source context:

Look after her a lot.
Okay.
Any questions?
Have we got her report?

Source sentence:

Yes, it's in the infirmary already.

Source context you highlighted:

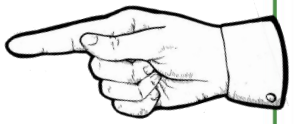
Reset Highlights

Source sentence you highlighted:

Reset Highlights

Target context:

Dorlotez-la.
D'accord.
Vous avez des questions?
On dispose de son rapport?



Target sentence:

Oui, ____ est à l'infirmerie.

- (•) **il**
- ( ) **elle**

How confident are you?

Not at all | Somewhat | Very

Target context you highlighted:

- rapport

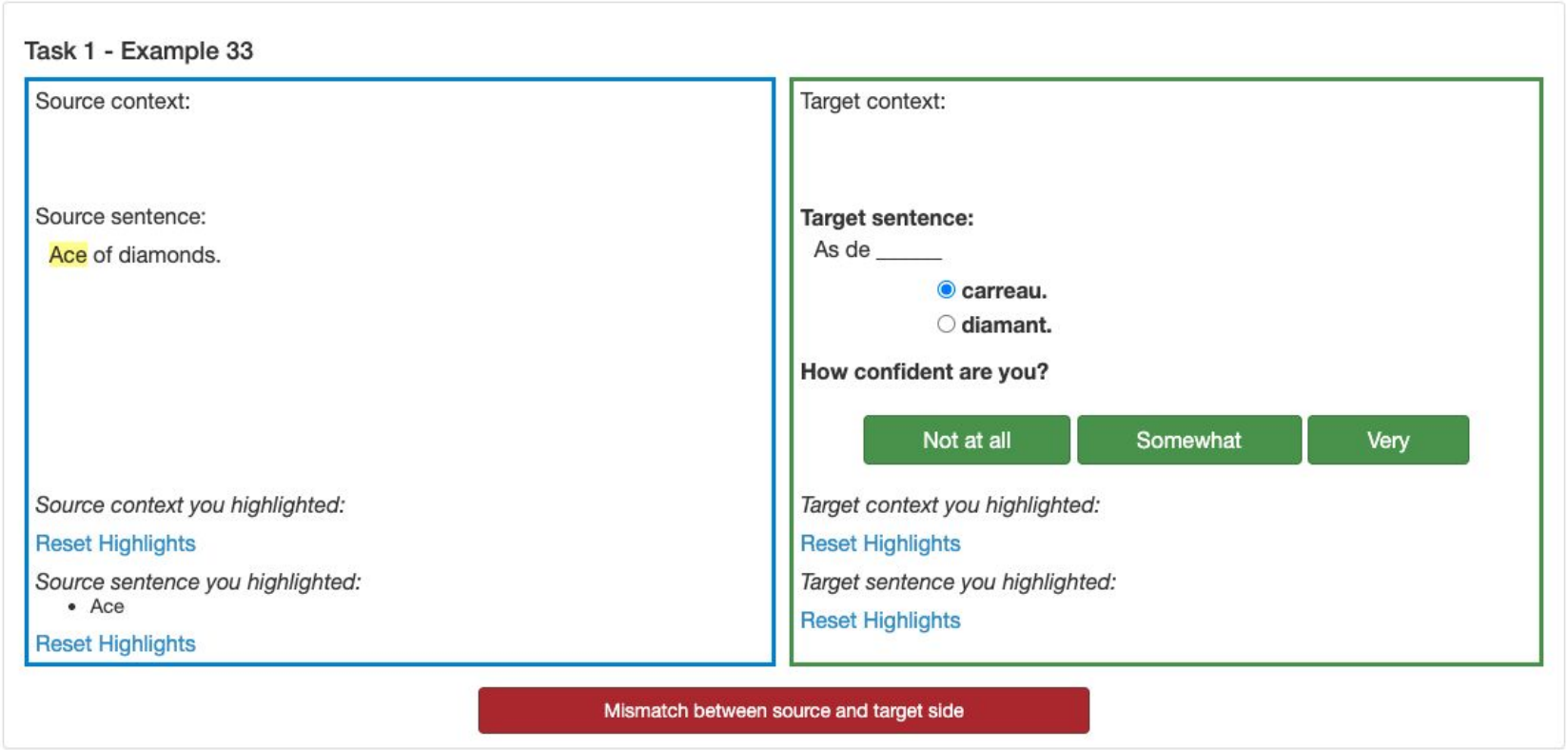
Reset Highlights

Target sentence you highlighted:

Reset Highlights

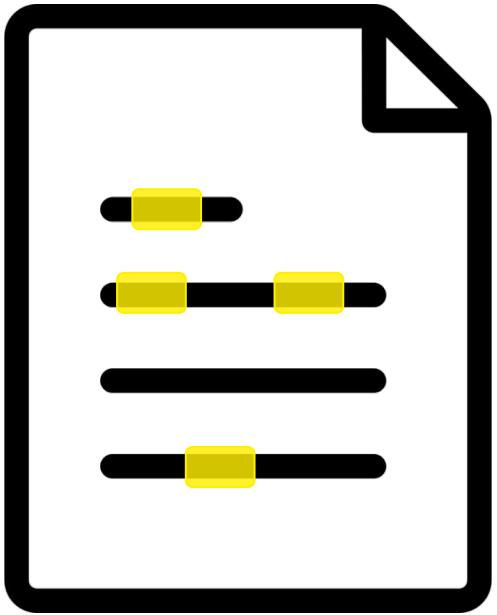
Mismatch between source and target side

User Study


Task 1 - Example 33
Source context:
Source sentence:
Ace of diamonds.
Source context you highlighted:
Reset Highlights
Source sentence you highlighted:
Ace
Reset Highlights
Target context:
Target sentence:
As de _____
carreau.
diamant.
How confident are you?
Not at all
Somewhat
Very
Target context you highlighted:
Reset Highlights
Target sentence you highlighted:
Reset Highlights
Mismatch between source and target side

What Context do Human Translators Use?

What Context do Human Translators Use?



What Context do Human Translators Use? (Pronoun Anaphora Resolution)

What Context do Human Translators Use? (Pronoun Anaphora Resolution)

What Context do Human Translators Use? (Pronoun Anaphora Resolution)

Have we got her report? Yes, it's in the infirmary already.
On dispose de son rapport? Oui, [**il** / ~~elle~~] est à l'infirmière.

What Context do Human Translators Use? (Pronoun Anaphora Resolution)

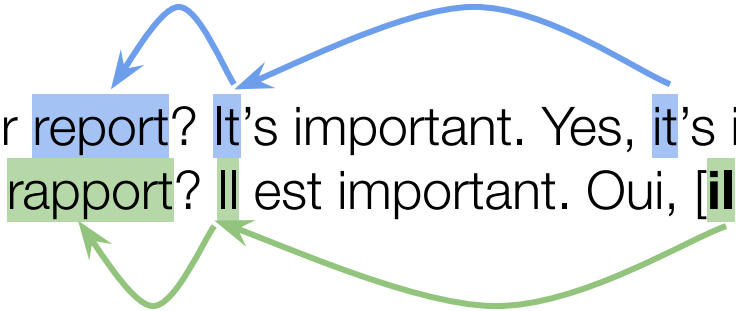
Have we got her report? Yes, it's in the infirmary already.
On dispose de son rapport? Oui, [**il** / ~~elle~~] est à l'infirmière.

What Context do Human Translators Use? (Pronoun Anaphora Resolution)

Have we got her report? It's important. Yes, it's in the infirmary already.

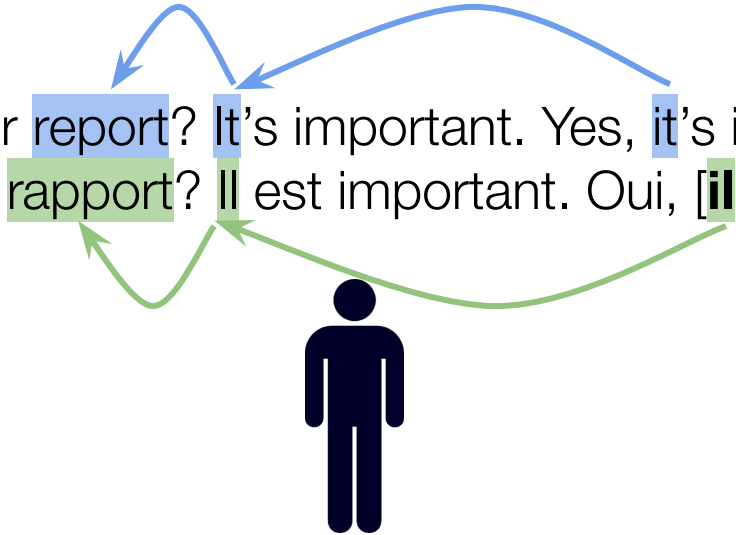
On dispose de son rapport? Il est important. Oui, [**il** / ~~elle~~] est à l'infirmière.

What Context do Human Translators Use? (Pronoun Anaphora Resolution)



Have we got her report? It's important. Yes, it's in the infirmary already.
On dispose de son rapport? Il est important. Oui, [**il** / ~~elle~~] est à l'infirmière.

What Context do Human Translators Use? (Pronoun Anaphora Resolution)



What Context do Human Translators Use? (Pronoun Anaphora Resolution)

Have we got her report? It's important. Yes, it's in the infirmary already.
On dispose de son rapport? Il est important. Oui, [**il** / ~~elle~~] est à l'infirmière.

What Context do Human Translators Use? (Word Sense Disambiguation)

What Context do Human Translators Use? (Word Sense Disambiguation)

What Context do Human Translators Use? (Word Sense Disambiguation)

Your charm is only exceeded by your frankness.

Ton [**charme** / ~~portebonheur~~] n’a d’égal que ta franchise.

What Context do Human Translators Use? (Word Sense Disambiguation)

Ton [**charme** / ~~portebonheur~~] n'a d'égal que ta franchise.

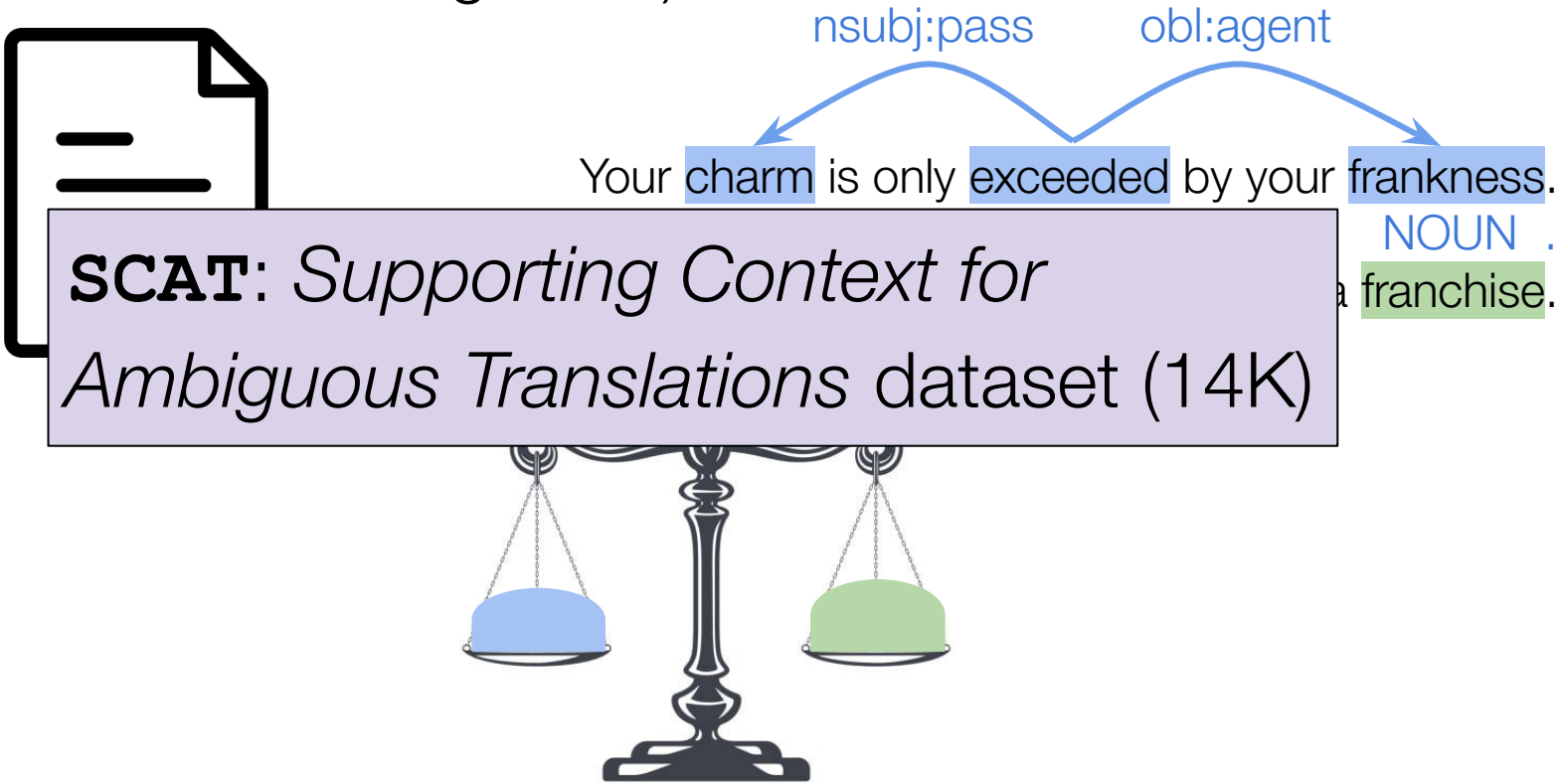
What Context do Human Translators Use? (Word Sense Disambiguation)

nsubj:pass obl:agent

Your charm is only exceeded by your frankness.
VERB NOUN .

Ton [**charme** / ~~portebonheur~~] n'a d'égal que ta franchise.

What Context do Human Translators Use? (Word Sense Disambiguation)

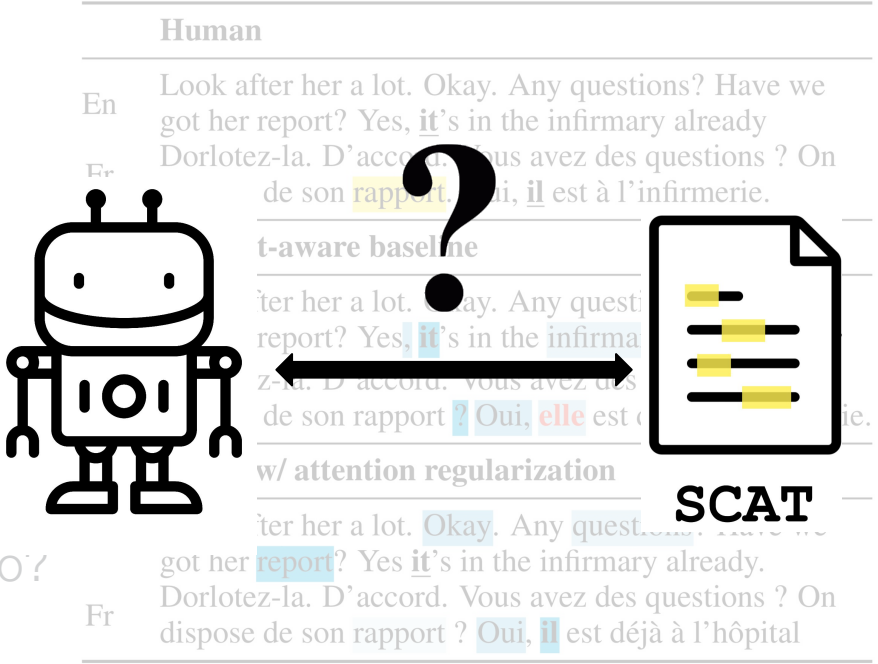


Outline

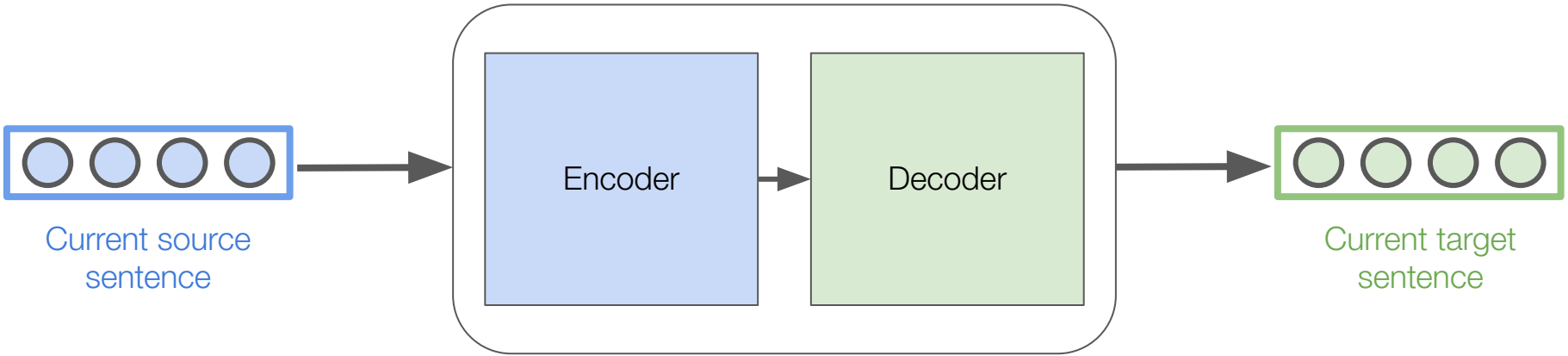
1. What context is useful during translation?

2. Are models paying attention to this context or not?

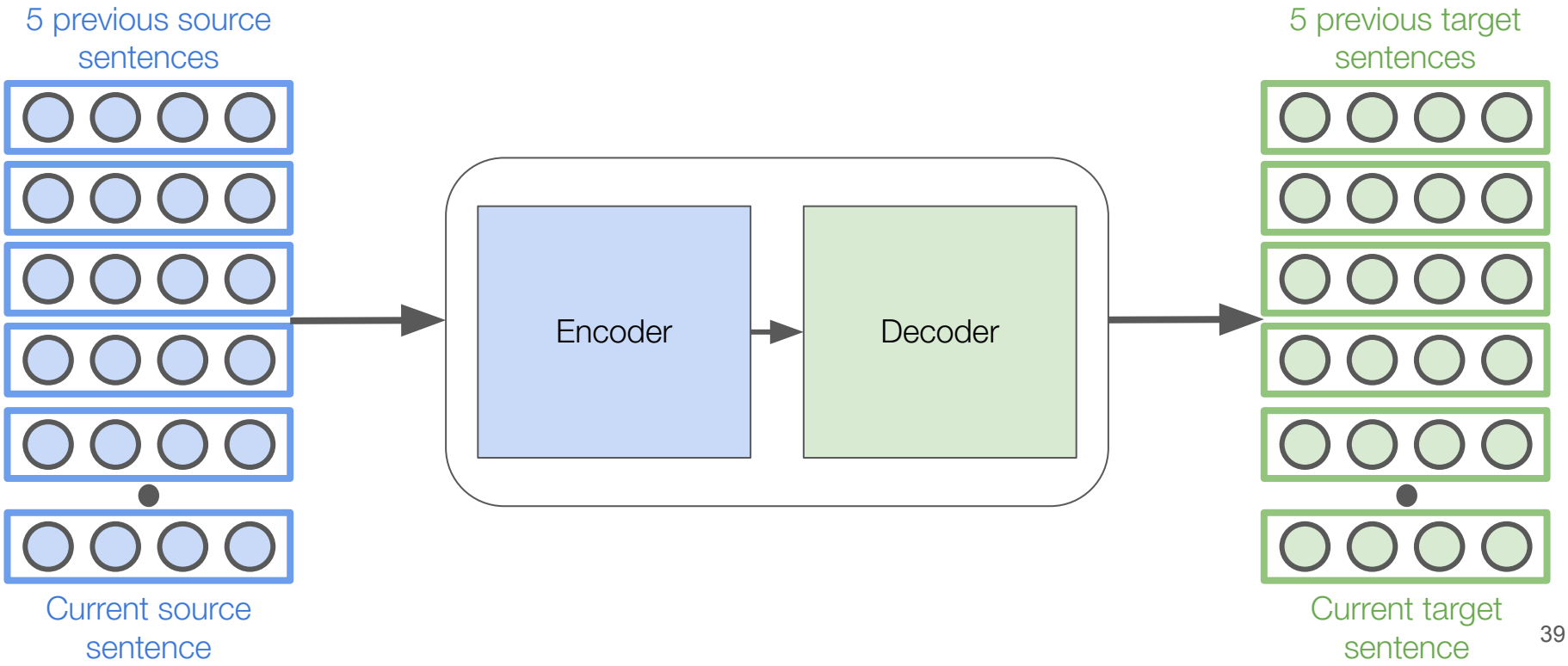
3. If not, can we encourage them to do so?



Model

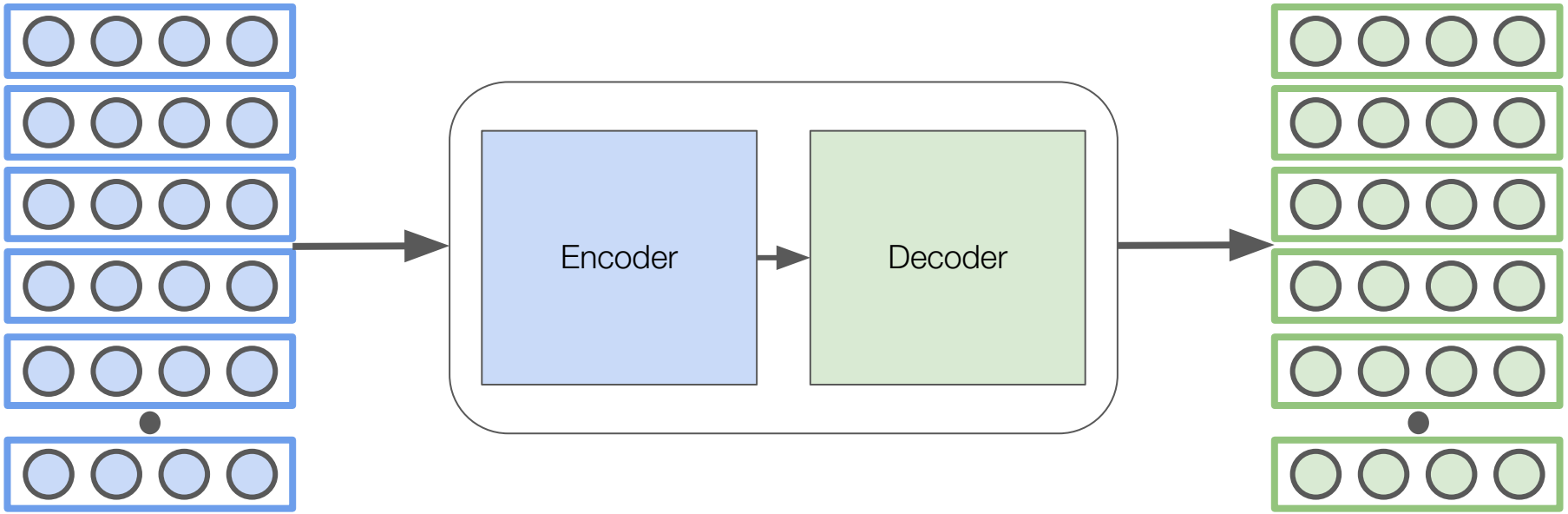

Encoder
Decoder
Current source sentence
Current target sentence

Model


5 previous source sentences
Encoder
Decoder
5 previous target sentences
Current source sentence
Current target sentence

Model

Open Subtitles



Quantifying Human-Model Alignment with **SCAT**

En Have we got her report?
Yes, it's in the infirmary already.

Fr On dispose de son rapport?
Oui, il est à l'infirmière.

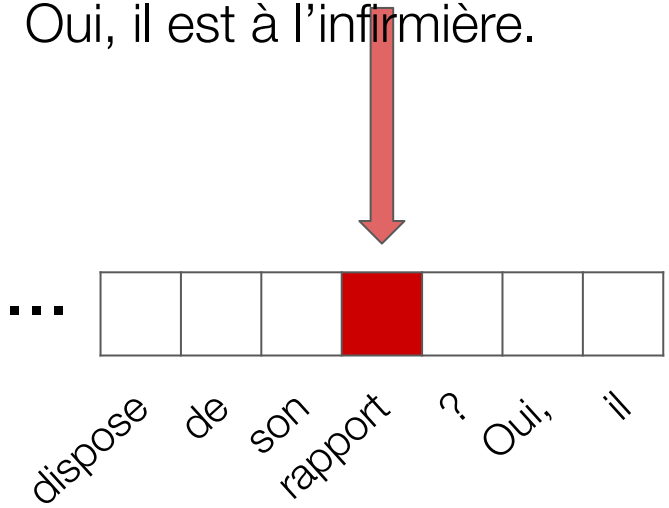
En Have we got her report?
Yes, it's in the infirmary already.

Fr On dispose de son rapport?
Oui, il est à l'infirmière.

Quantifying Human-Model Alignment with **SCAT**

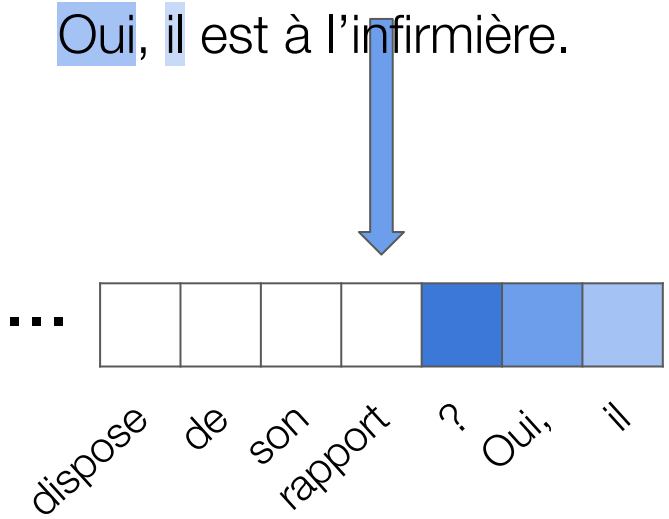
En Have we got her report?
Yes, it's in the infirmary already.

Fr On dispose de son rapport?
Oui, il est à l'infirmière.

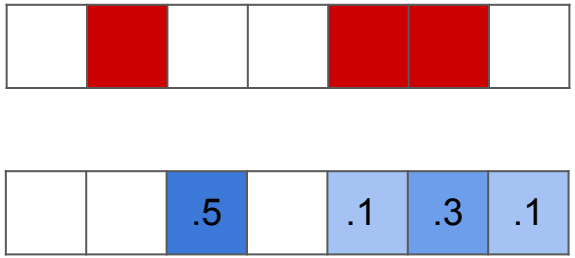


En Have we got her report?
Yes, it's in the infirmary already.

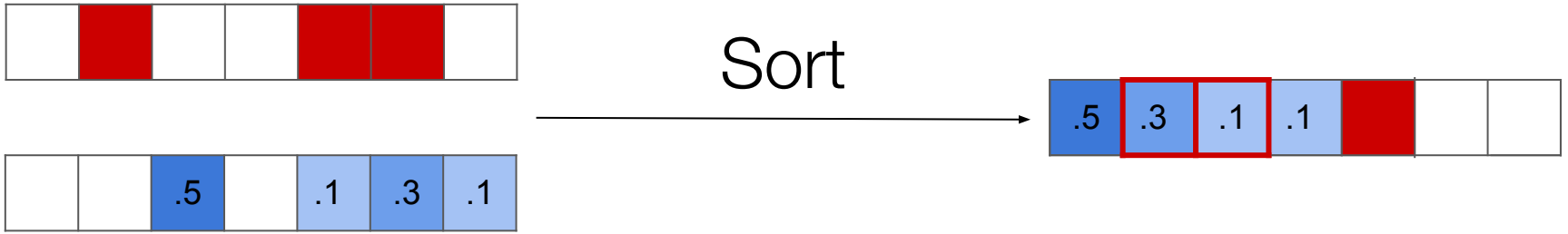
Fr On dispose de son rapport?
Oui, il est à l'infirmière.



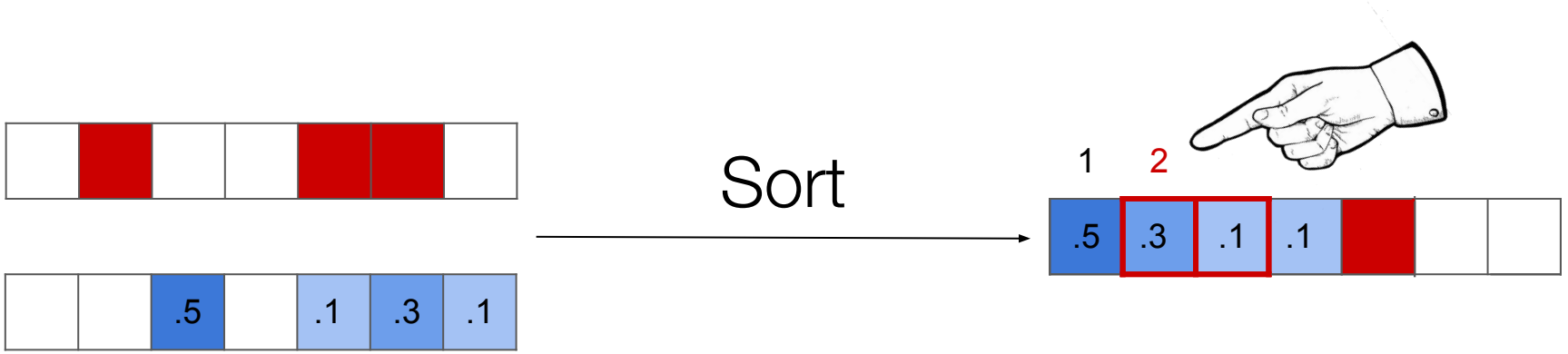
Quantifying Human-Model Alignment with **SCAT**



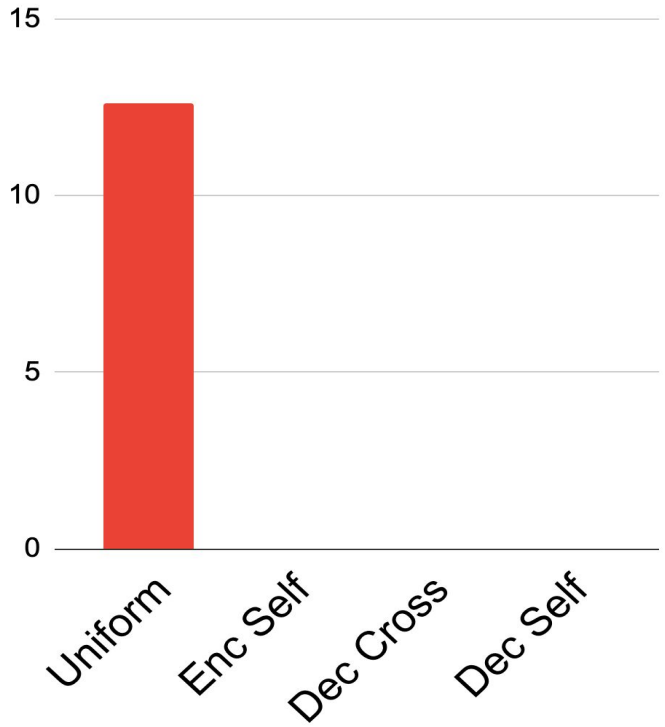
Quantifying Human-Model Alignment with **SCAT**



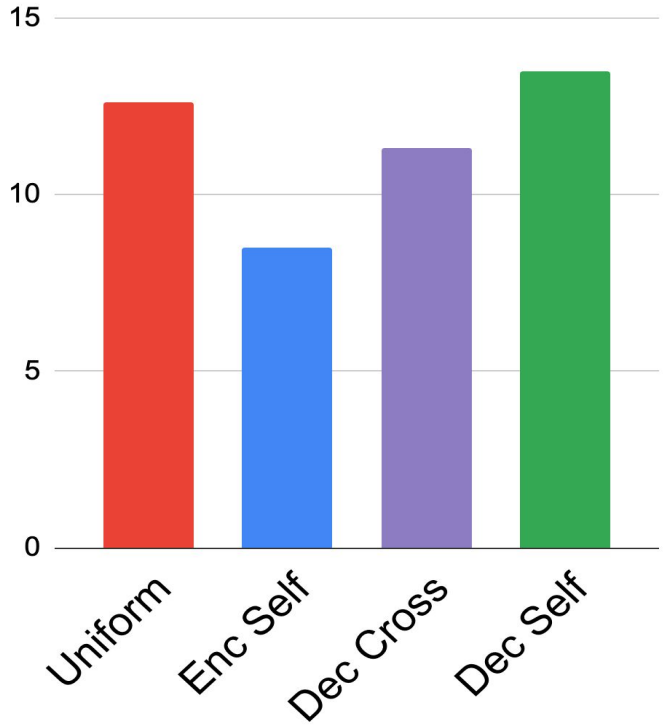
Quantifying Human-Model Alignment with **SCAT**



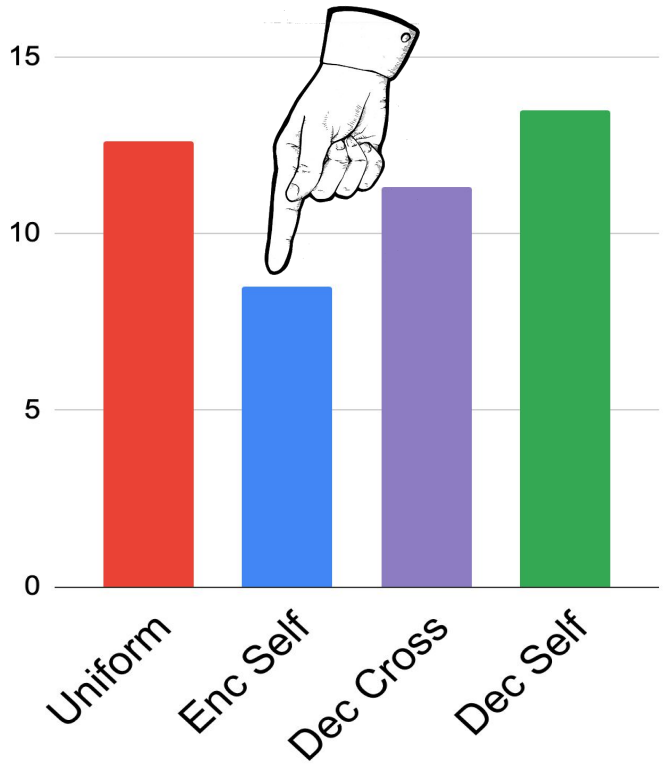
Alignment Results



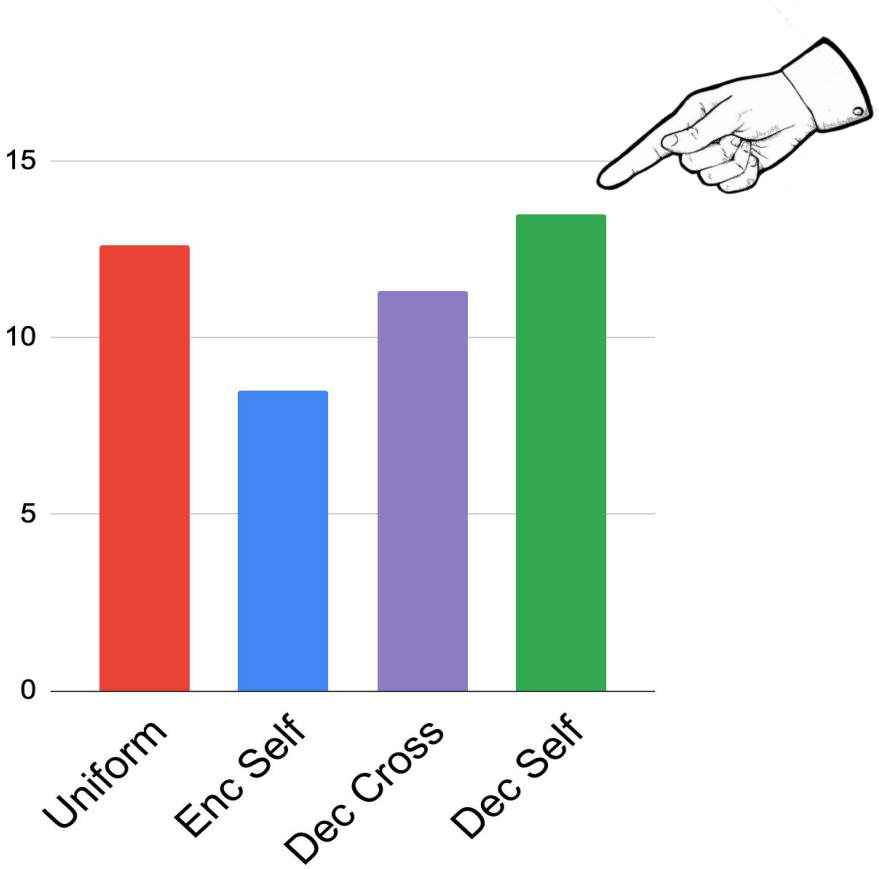
Alignment Results



Alignment Results



Alignment Results

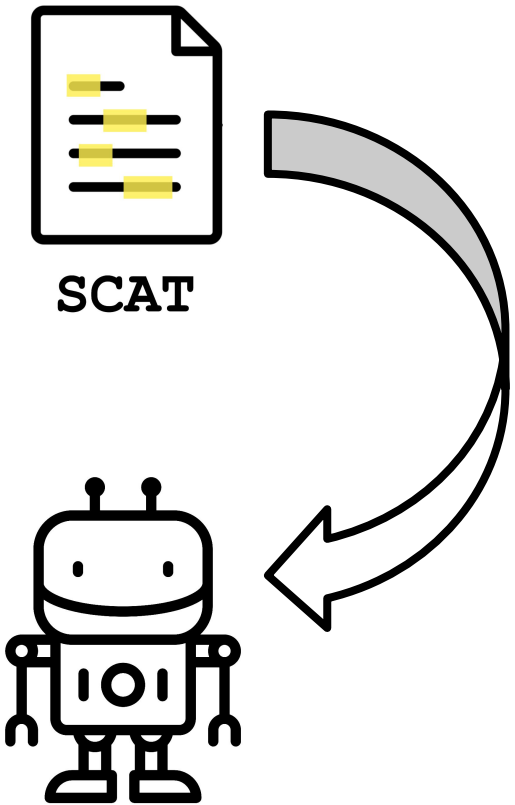


Outline

1. What context is useful during translation?

2. Are models paying attention to this context or not?

3. If not, can we encourage them to do so?



Attention Regularization

$$\mathcal{L}_{NLL}(\theta) = -\sum_{j=1}^{m} \log p_\theta(y_j | x, y_{i<j})$$

OpenSubtitles18

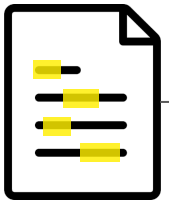
Context-aware MT Model

Attention Regularization

$$\mathcal{L}_{NLL}(\theta) = -\sum_{j=1}^{m} \log p_\theta(y_j | x, y_{i<j})$$

OpenSubtitles18

Context-aware MT Model



SCAT

$$\mathcal{R}(\theta) = -\lambda \mathrm{KL}(\alpha_{\text{human-norm}} || \alpha_{\text{model}}(\theta))$$

Evaluation

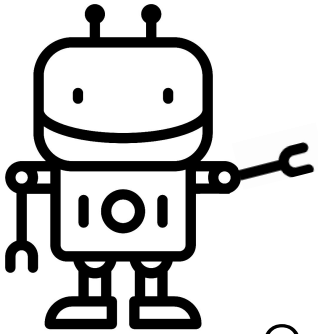
- BLEU
- COMET

Evaluation

- BLEU
- COMET
- Pronouns F-measure

Evaluation

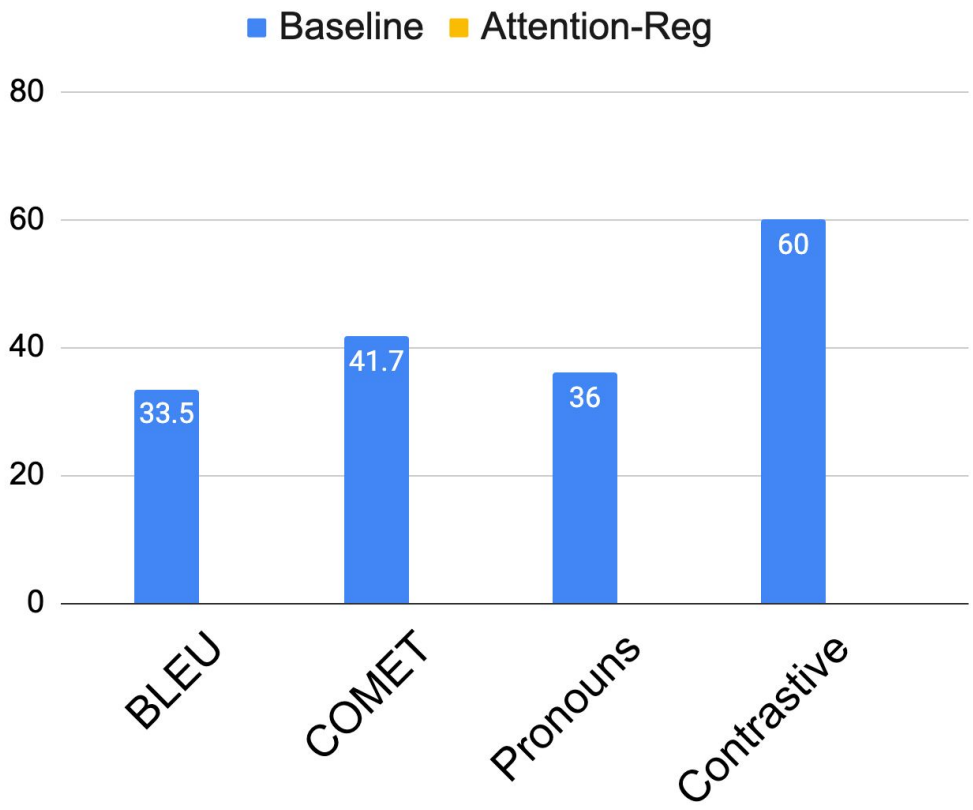
- BLEU
- COMET
- Pronouns F-measure
- Contrastive Evaluation



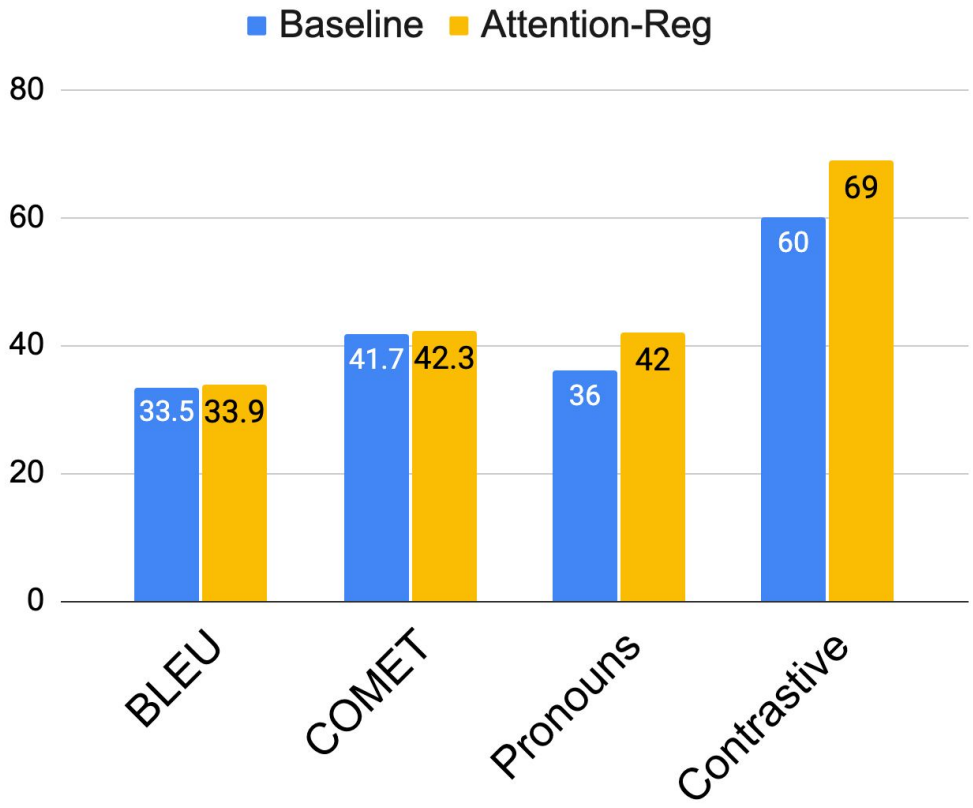
Oui, **il** est déjà à l'infirmerie.

Oui, **elle** est déjà à l'infirmerie.

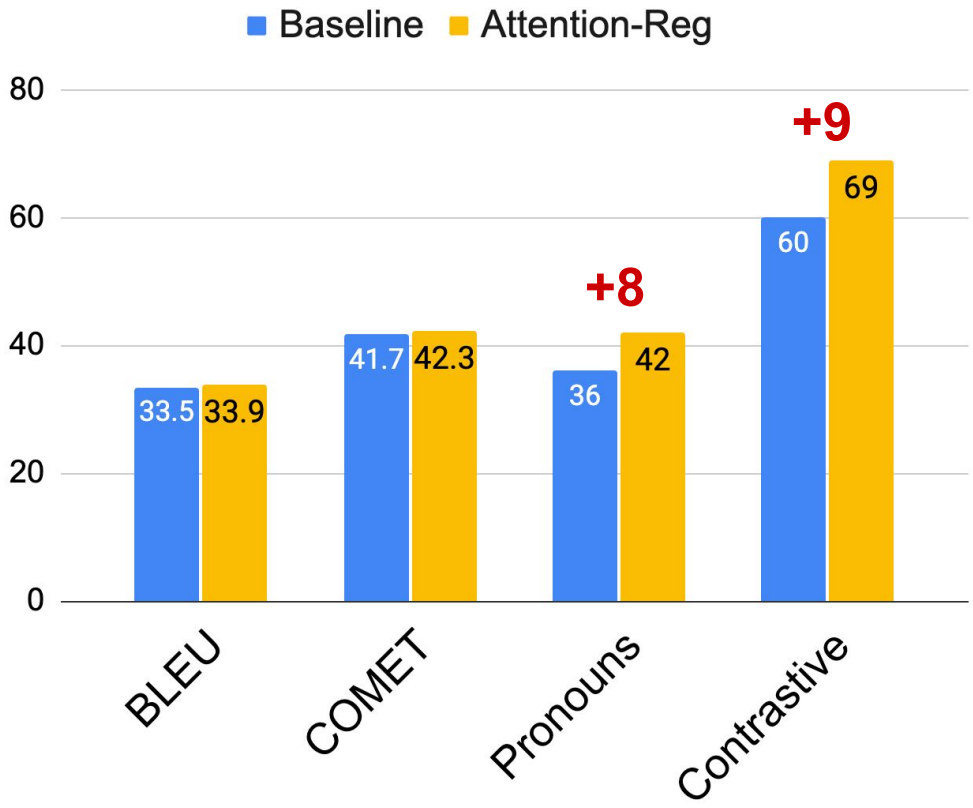
Results


Baseline
Attention-Reg
80
60
40
20
0
33.5
41.7
36
60
BLEU
COMET
Pronouns
Contrastive

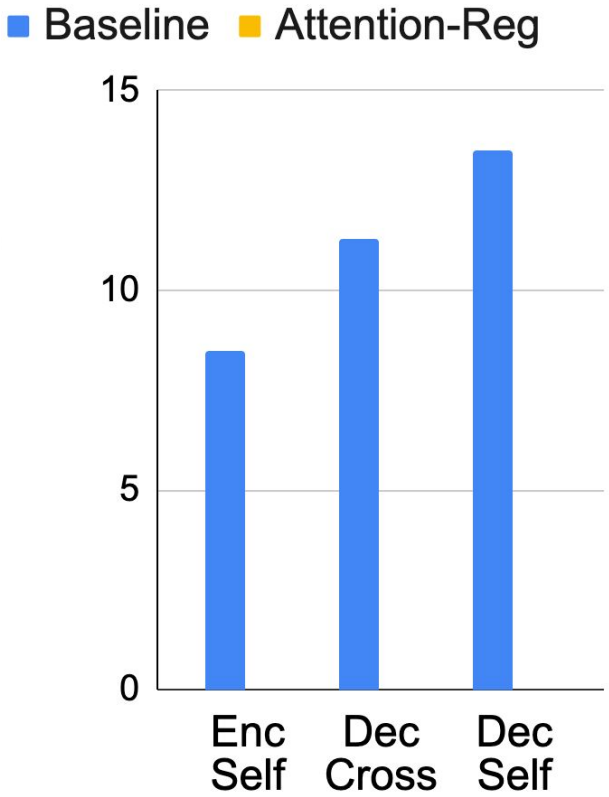
Results


Baseline
Attention-Reg
80
60
40
20
0
33.5
33.9
41.7
42.3
36
42
60
69
BLEU
COMET
Pronouns
Contrastive

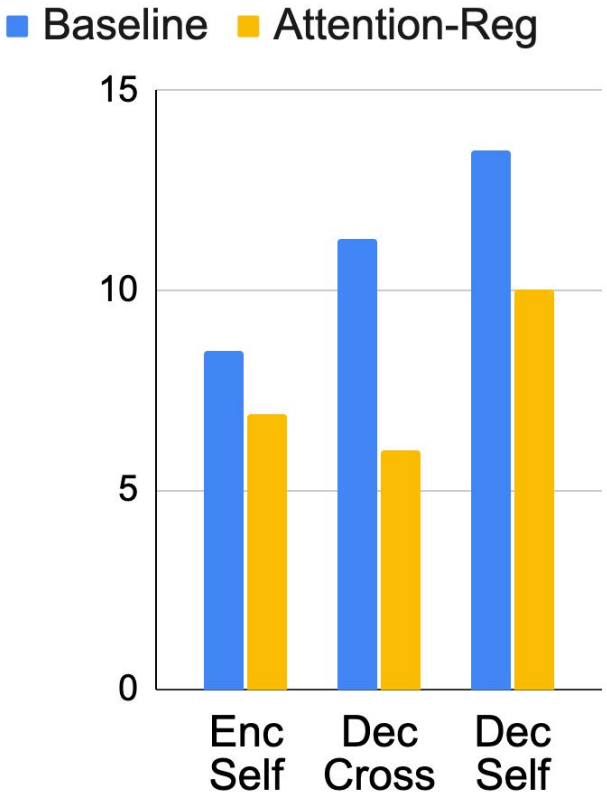
Results


Baseline
Attention-Reg
80
60
40
20
0
33.5
33.9
41.7
42.3
36
42
60
69
+8
+9
BLEU
COMET
Pronouns
Contrastive

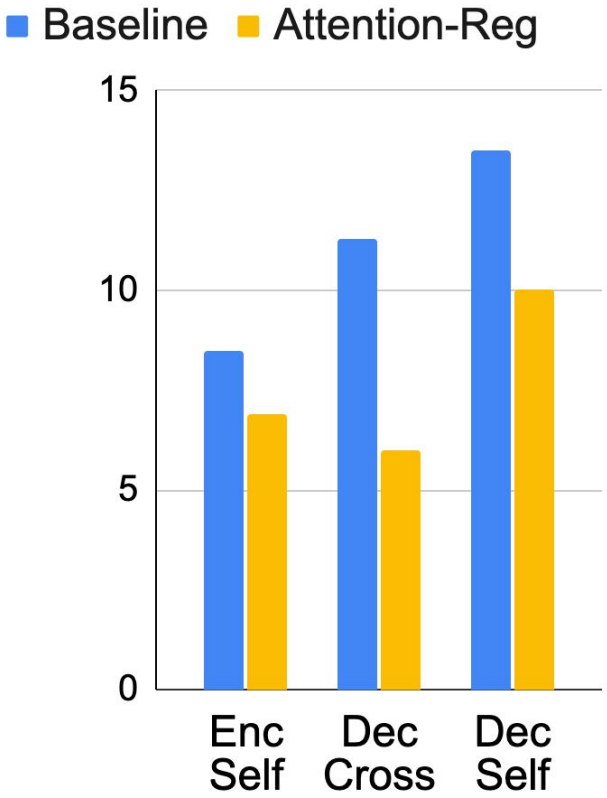
Results



Results



Results



Have we got her report?
Yes, it's in the infirmary already.

Baseline

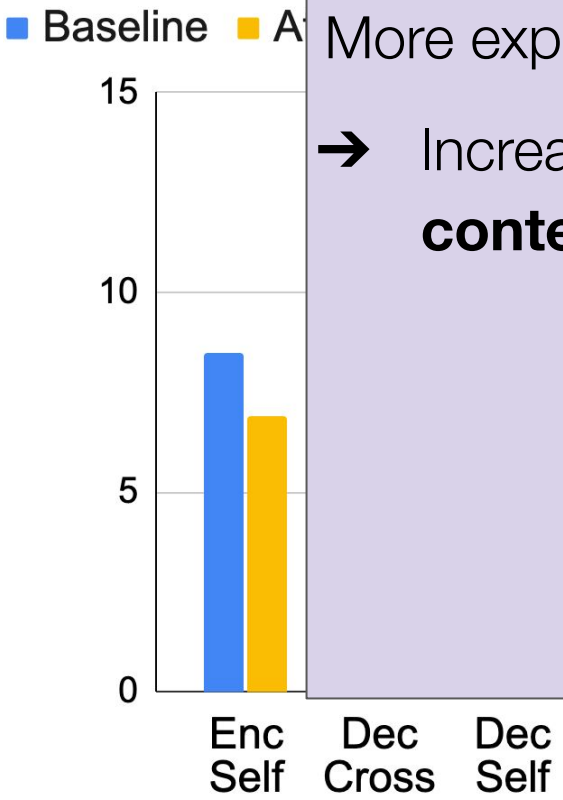
On dispose de son rapport?
Oui, **elle** est déjà à l'infirmerie.

Have we got her report?
Yes, it's in the infirmary already.

Attention-Reg

On dispose de son rapport?
Oui, **il** est déjà à l'infirmerie.

Results

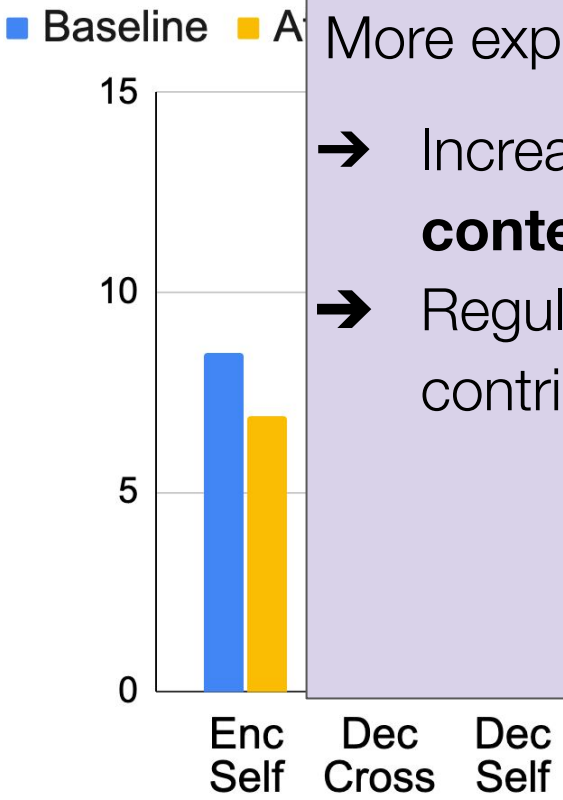


More experiments & results in paper:

➔ Increased usage of **supporting context**

Have we got her report?
mary already.
son rapport?
à l'infirmerie.
ot her report?
rmary already.
de son rapport?
Oui, il est déjà à l'infirmerie.

Results



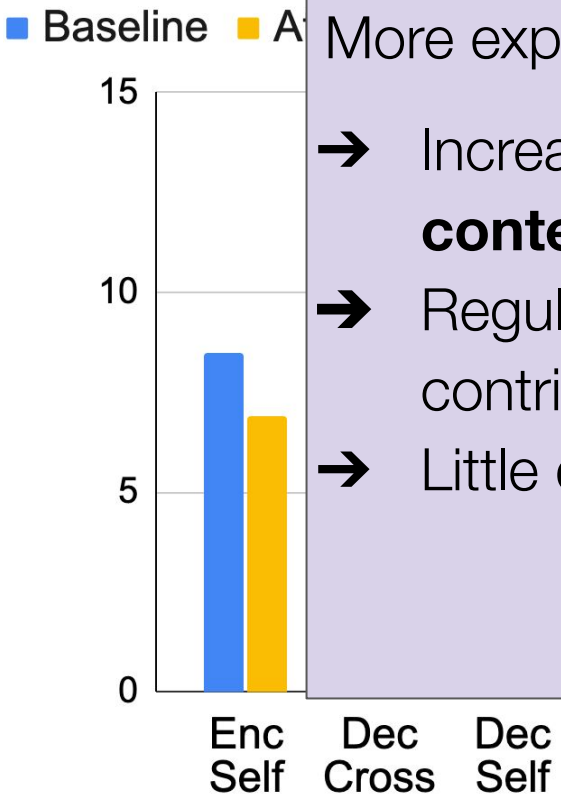
More experiments & results in paper:

➔ Increased usage of **supporting context**

➔ Regularizing **encoder self-attention** contributes the most

Have we got her report?
mary already.
son rapport?
à l'infirmerie.
ot her report?
rmary already.
de son rapport?
Oui, il est déjà à l'infirmerie.

Results

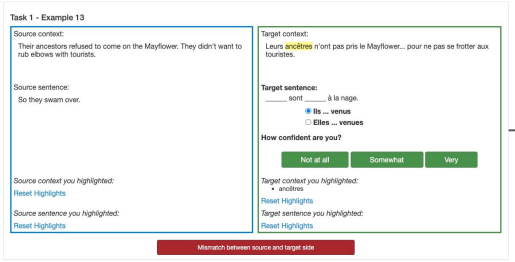


More experiments & results in paper:

- ➔ Increased usage of **supporting context**
- ➔ Regularizing **encoder self-attention** contributes the most
- ➔ Little difference in **WSD** performance

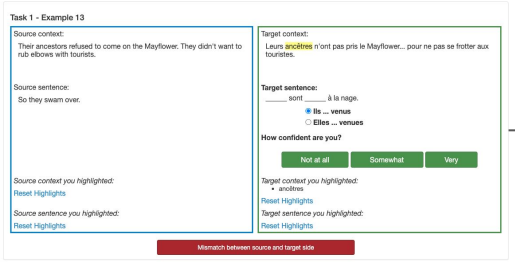
Have we got her report?
mary already.
son rapport?
à l'infirmerie.
ot her report?
rmary already.
de son rapport?
Oui, il est déjà à l'infirmerie.

Summary



14k SCAT

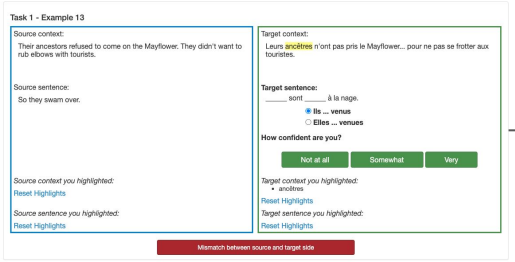
Summary



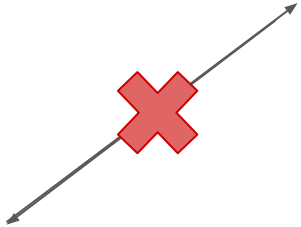
14k SCAT

Baseline Model

Summary

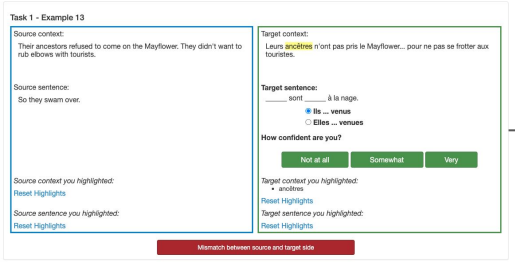


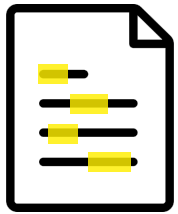
14k SCAT



Baseline Model

Summary





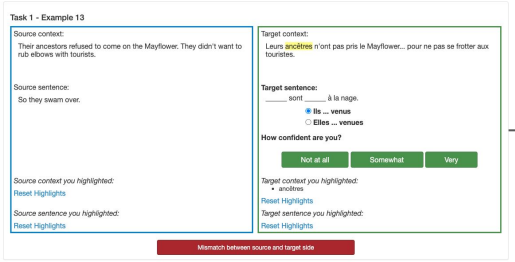
14k SCAT

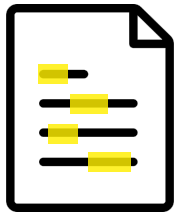
$$\mathcal{R}(\theta)$$

Regularized Model

Baseline Model

Summary





14k SCAT

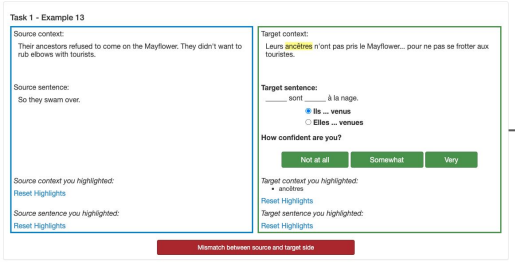
$\mathcal{R}(\theta)$

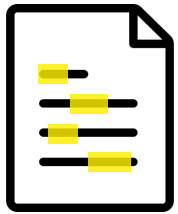
Regularized Model

Baseline Model

➔ Better alignment, context usage, translation performance

Summary





14k SCAT

$\mathcal{R}(\theta)$

Regularized Model

Baseline Model

- ➔ Better alignment, context usage, translation performance
- ➔ Code & Data: **github.com/neulab/contextual-mt**