Extending Neural Machine Translation to Documents and Signed Languages

Kayo Yin

Talk @ University of Pittsburgh

October 7 2021

Carnegie Mellon University
Language Technologies Institute

Introduction

➔ Neural Machine Translation is the state-of-the-art in automated translation

Introduction

- ➔ Neural Machine Translation is the state-of-the-art in automated translation

- ➔ However, they are usually limited to **sentence-level** translations

Introduction

➔ Neural Machine Translation is the state-of-the-art in automated translation

➔ However, they are usually limited to **sentence-level** translations

➔ Current NLP systems also cannot process **signed languages**

Today's Agenda

➔ Do context-aware machine translation models **pay the right attention**?

Today’s Agenda

➔ Do context-aware machine translation models **pay the right attention**?

➔ When does translation require **context**?

Today’s Agenda

- ➔ Do context-aware machine translation models **pay the right attention**?
- ➔ When does translation require **context**?
- ➔ How do we resolve **coreference** in **signed languages**?

Do Context-Aware Translation Models Pay the Right Attention?

Kayo Yin, Patrick Fernandes, Danish Pruthi, Aditi Chaudhary
André F.T. Martins, Graham Neubig

(ACL 2021)

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

- Many approaches have been proposed for **context-aware** machine translation

Context-Aware NMT

- Many approaches have been proposed for **context-aware** machine translation
  - Concatenation, Multi-Encoder, Cache-Based, Hierarchical…

Context-Aware NMT

- Many approaches have been proposed for **context-aware** machine translation
  - Concatenation, Multi-Encoder, Cache-Based, Hierarchical…

- Most of these approaches perform poorly on document-level translation

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.

- [x] **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:

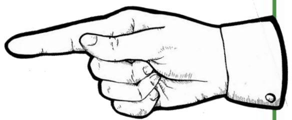
Reset Highlights

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.

- (selected) **il**
- **elle**

How confident are you?

Not at all | Somewhat | Very

Target context you highlighted:

- rapport

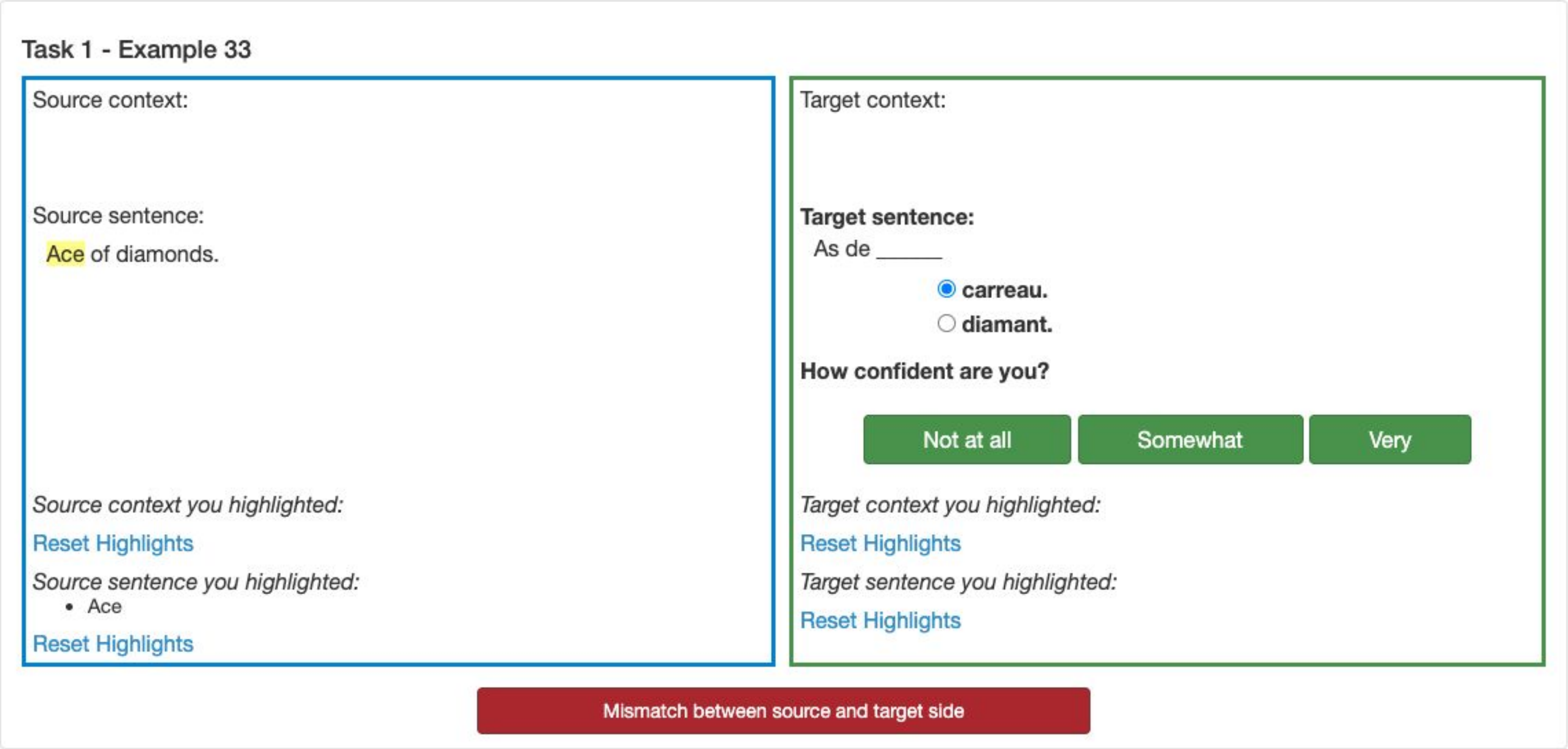
Reset Highlights

Target sentence you highlighted:

Reset Highlights

Mismatch between source and target side

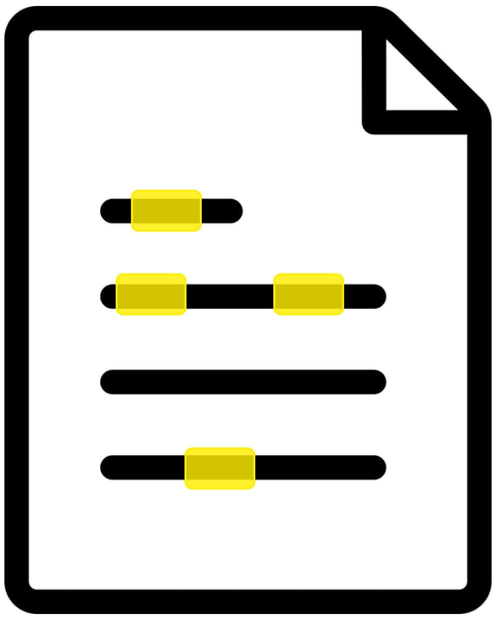
User Study



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)

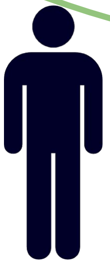
Have we got her report? It’s important. Yes, it’s in the infirmary already.

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

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

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.

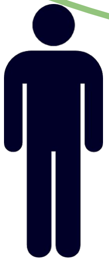


What Context do Human Translators Use? (Pronoun Anaphora Resolution)

?

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?
(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)

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)

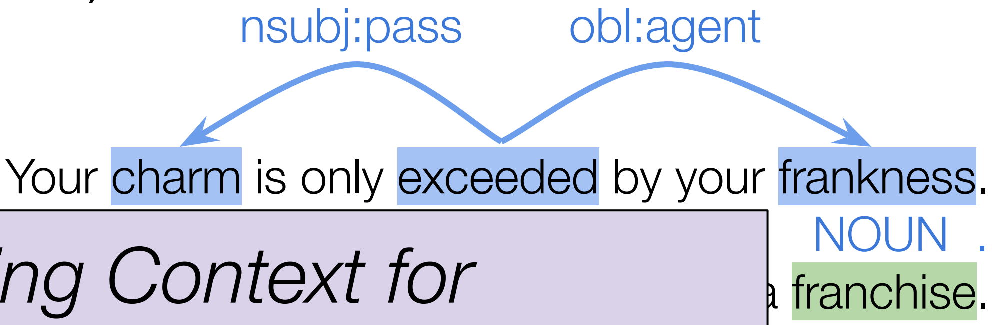
nsubj:pass obl:agent

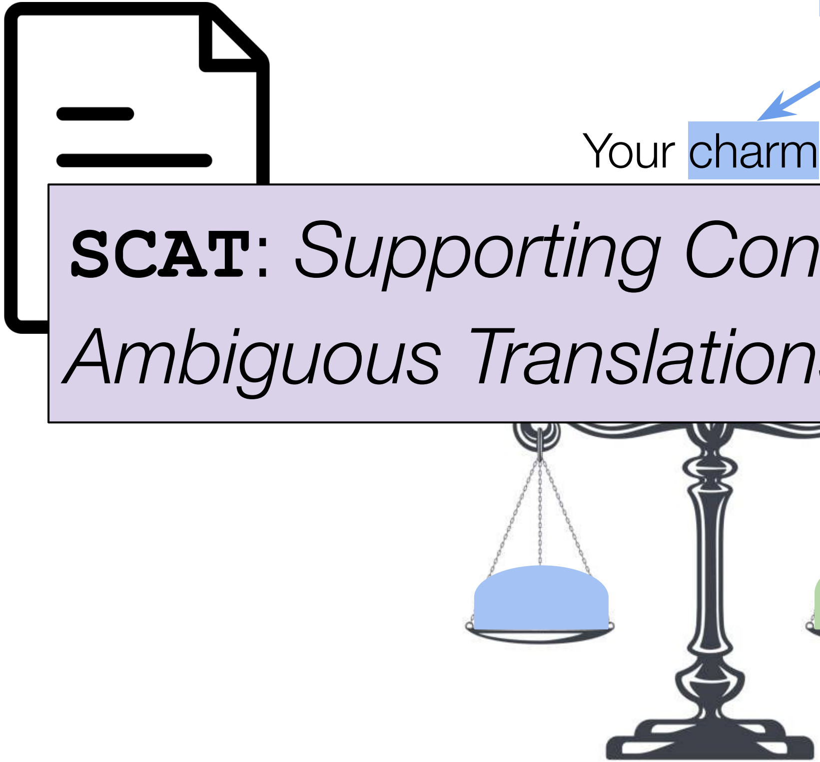
Your charm is only exceeded by your frankness.

VERB NOUN .

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

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





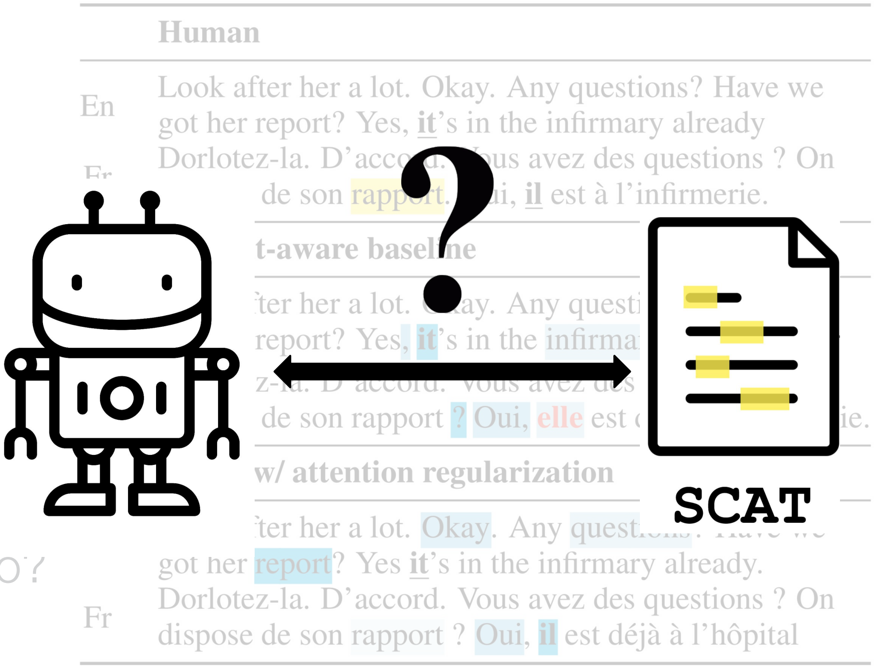
SCAT: *Supporting Context for Ambiguous Translations* dataset (14K)

Outline

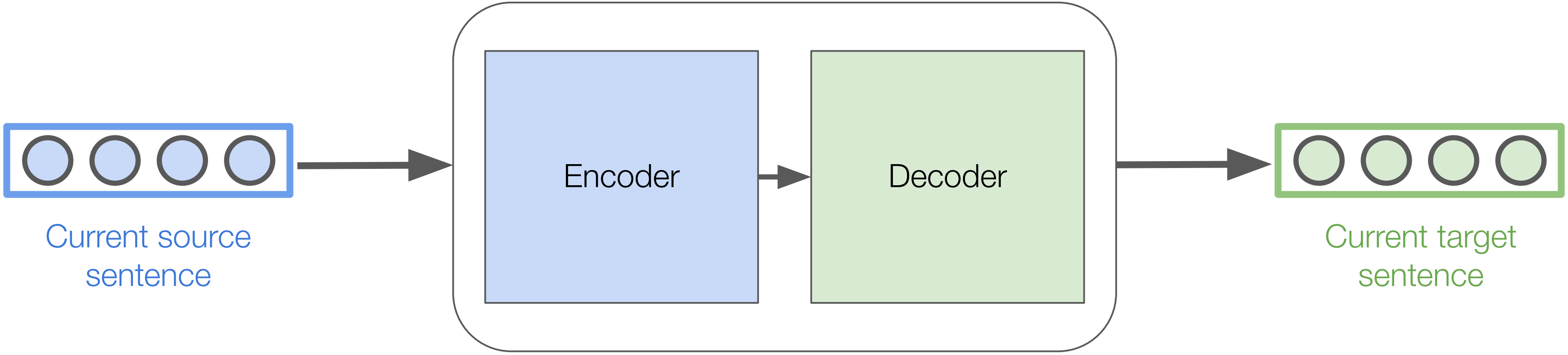
1. What context is useful during translation?

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

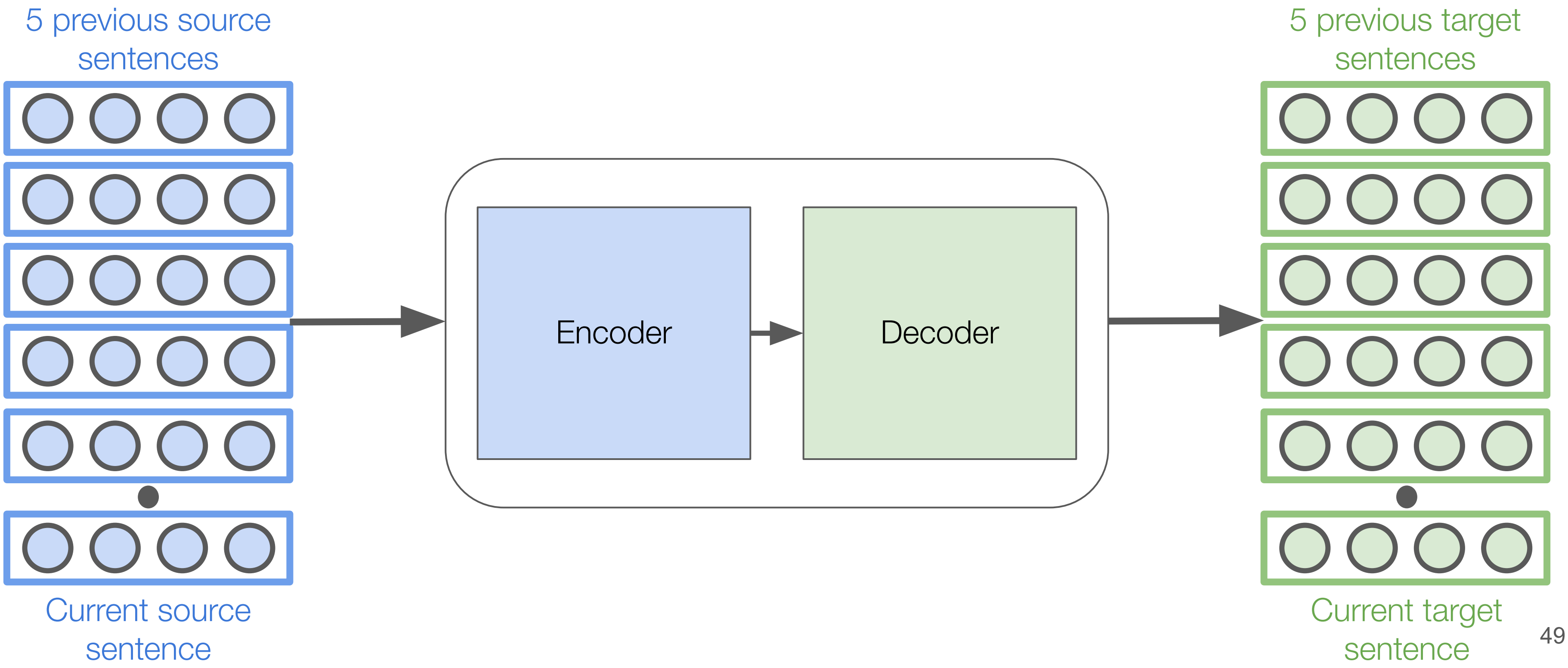
3. If not, can we encourage them to do so?



Model

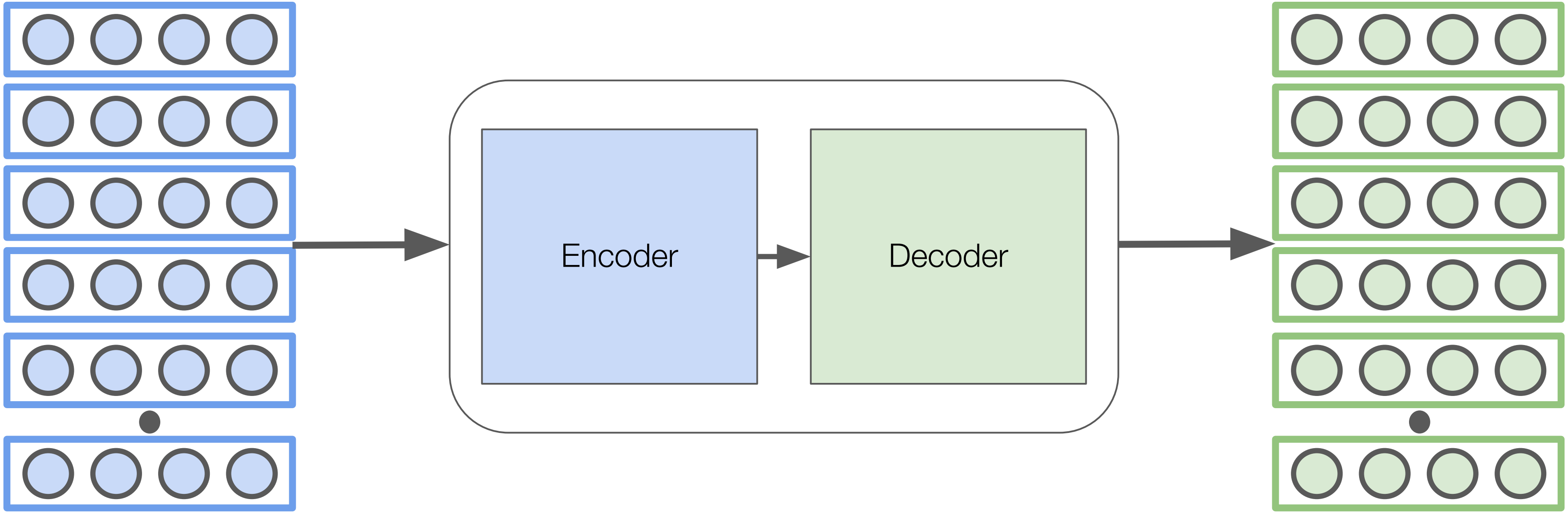

Encoder
Decoder
Current source sentence
Current target sentence

Model


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

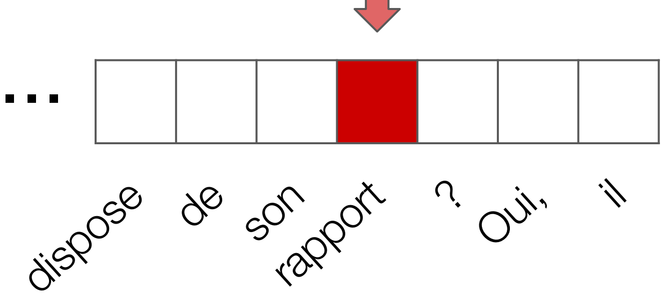
Fr

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

Quantifying Human-Model Alignment with **SCAT**

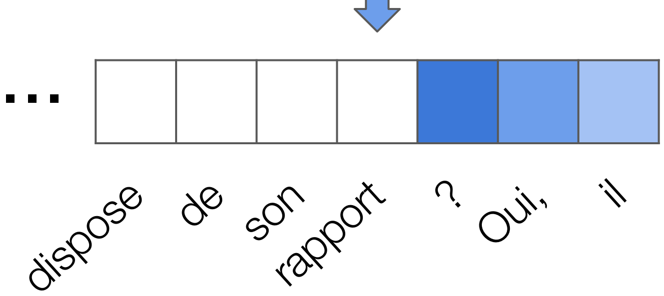
En
Have we got her report?
Yes, it's in the infirmary already.

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

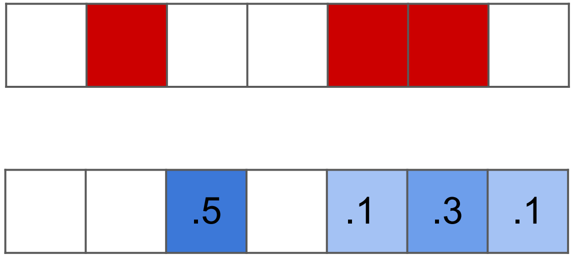


En
Have we got her report?
Yes, it's in the infirmary already.

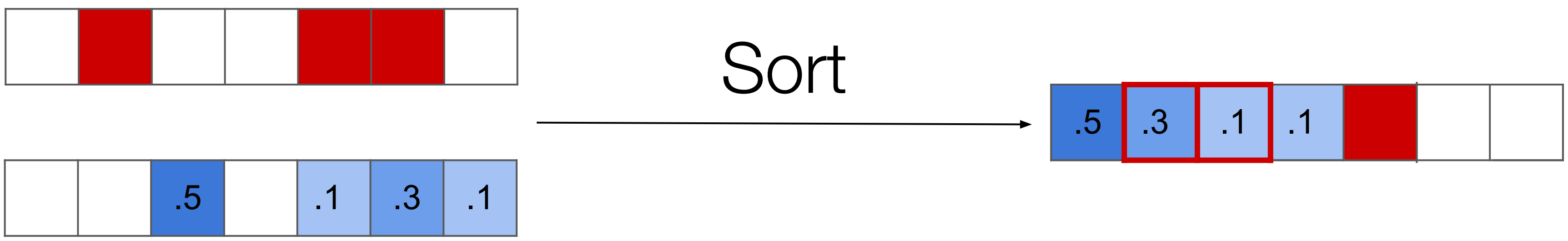
Fr
On dispose de son rapport?
Oui, il est à l'infirmière.



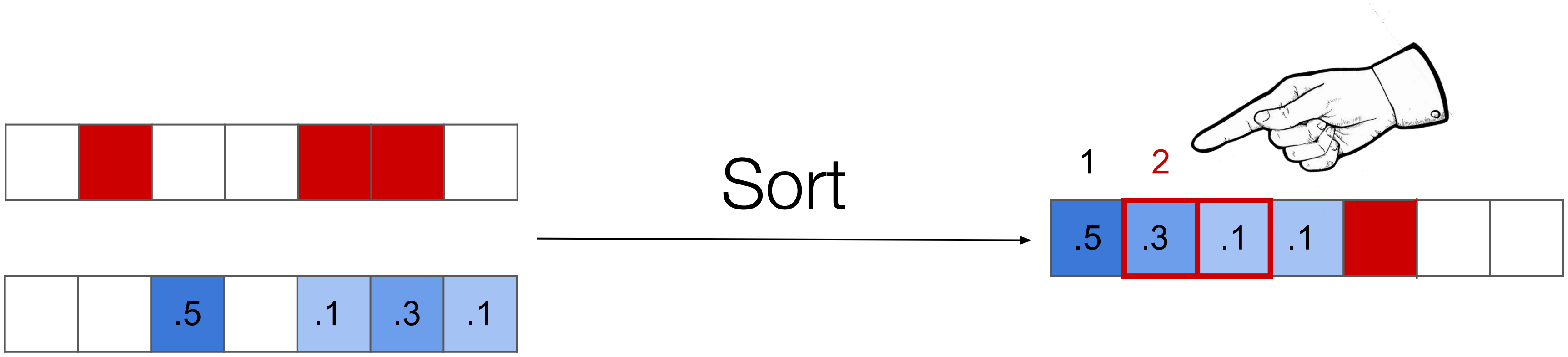
Quantifying Human-Model Alignment with **SCAT**



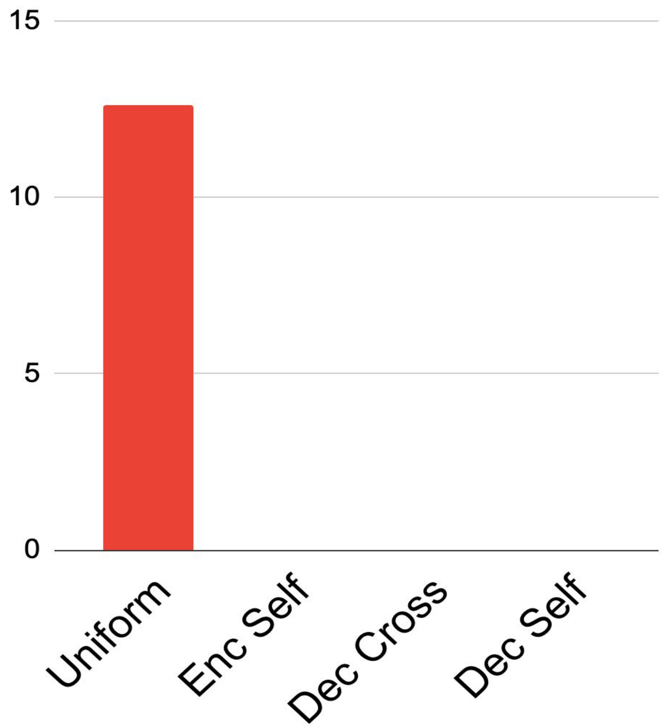
Quantifying Human-Model Alignment with **SCAT**



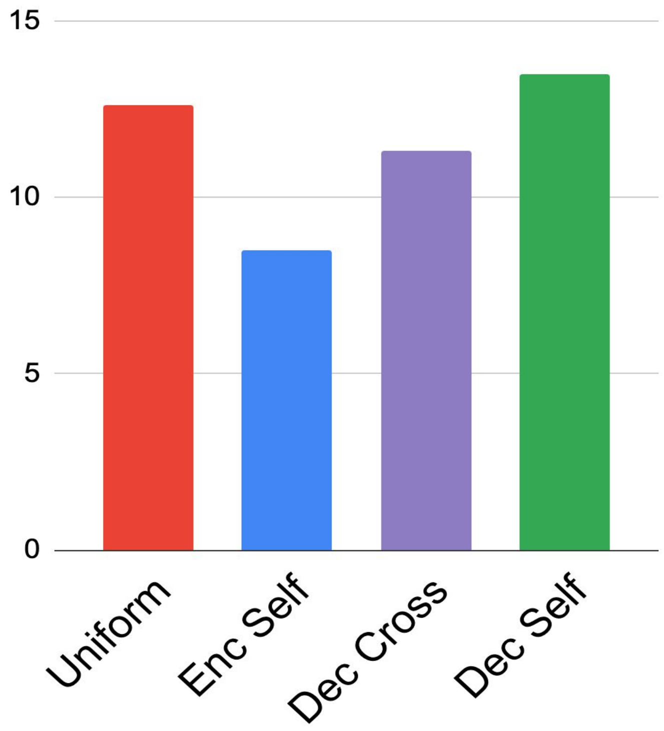
Quantifying Human-Model Alignment with **SCAT**



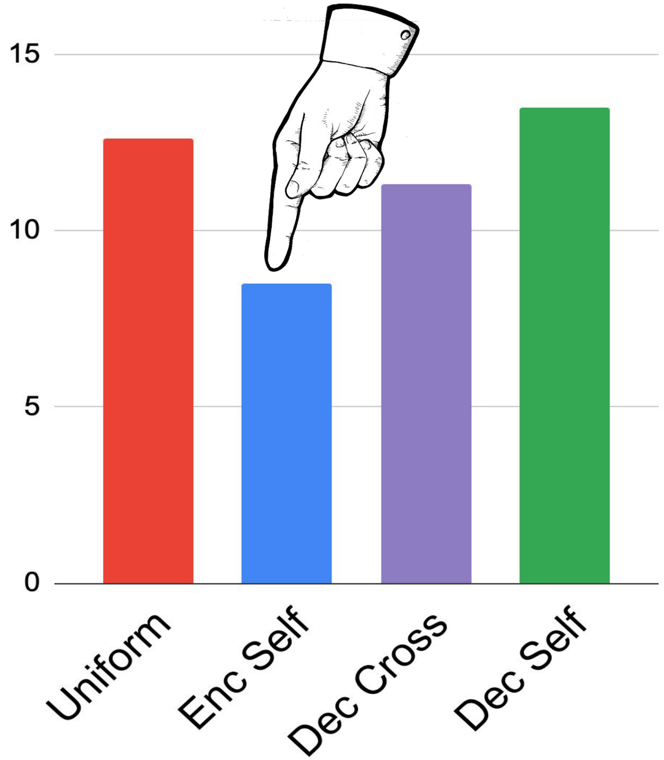
Alignment Results



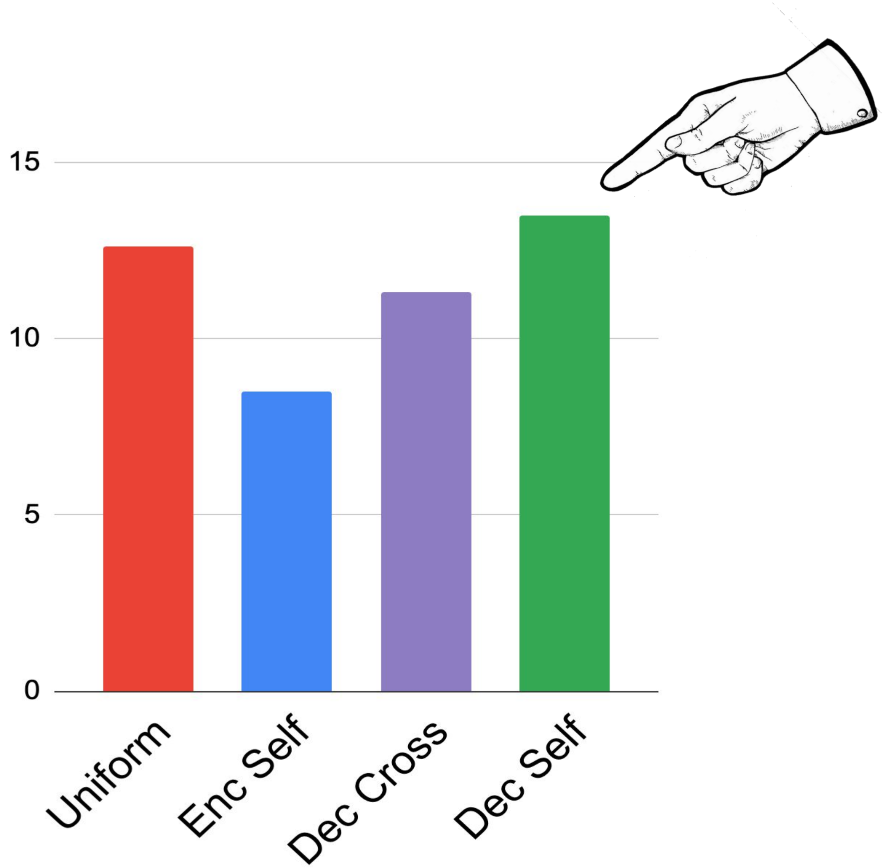
Alignment Results



Alignment Results



Alignment Results

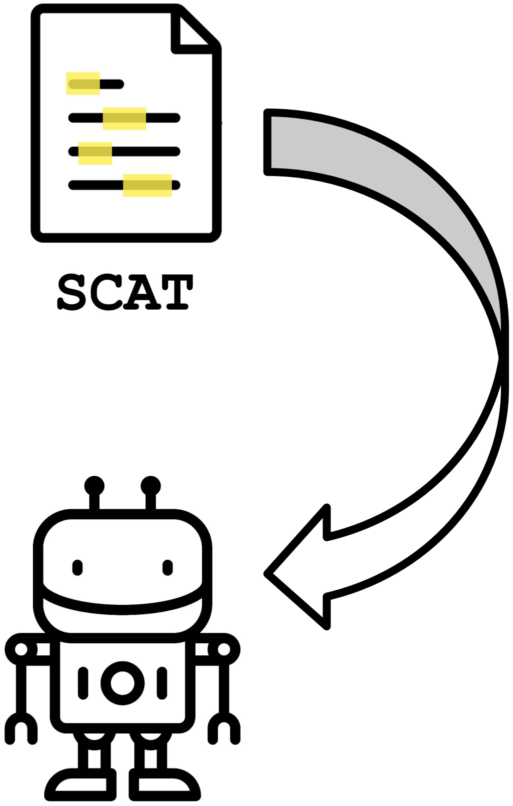


Outline

1. What context is useful during translation?

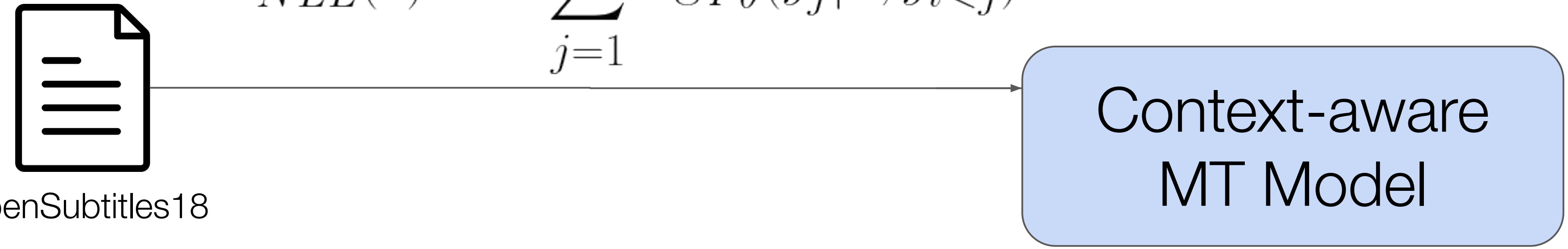
2. Are models paying attention to this context or not?

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



Attention Regularization

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



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

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

Evaluation

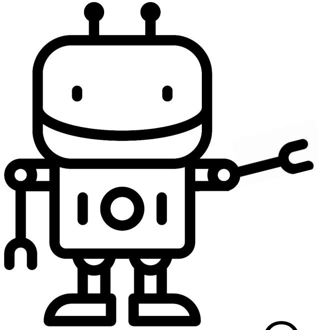
- BLEU
- COMET

Evaluation

- BLEU
- COMET
- Pronouns F-measure

Evaluation

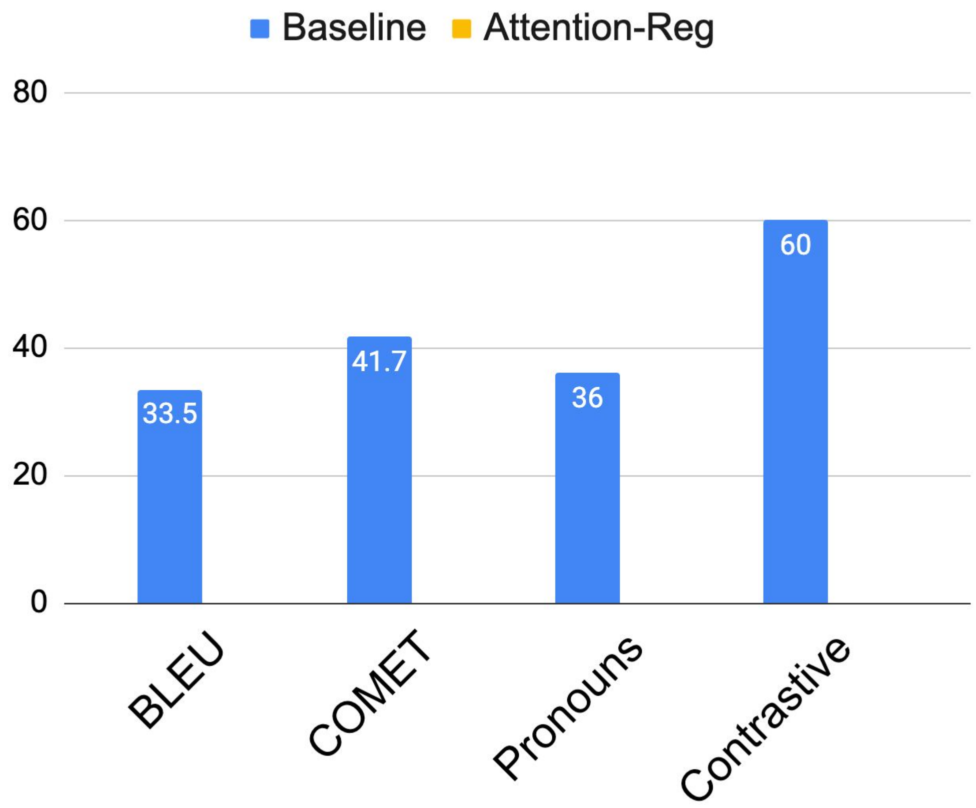
- BLEU
- COMET
- Pronouns F-measure
- Contrastive Evaluation



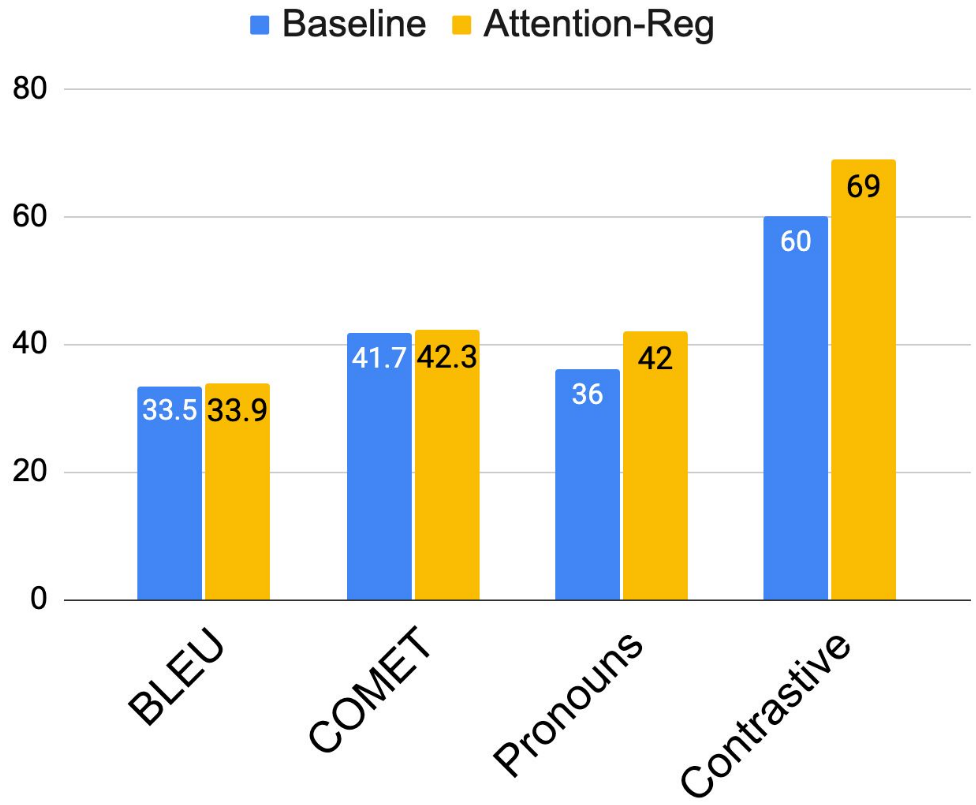
Oui, **il** est déjà à l’infirmerie.

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

Results

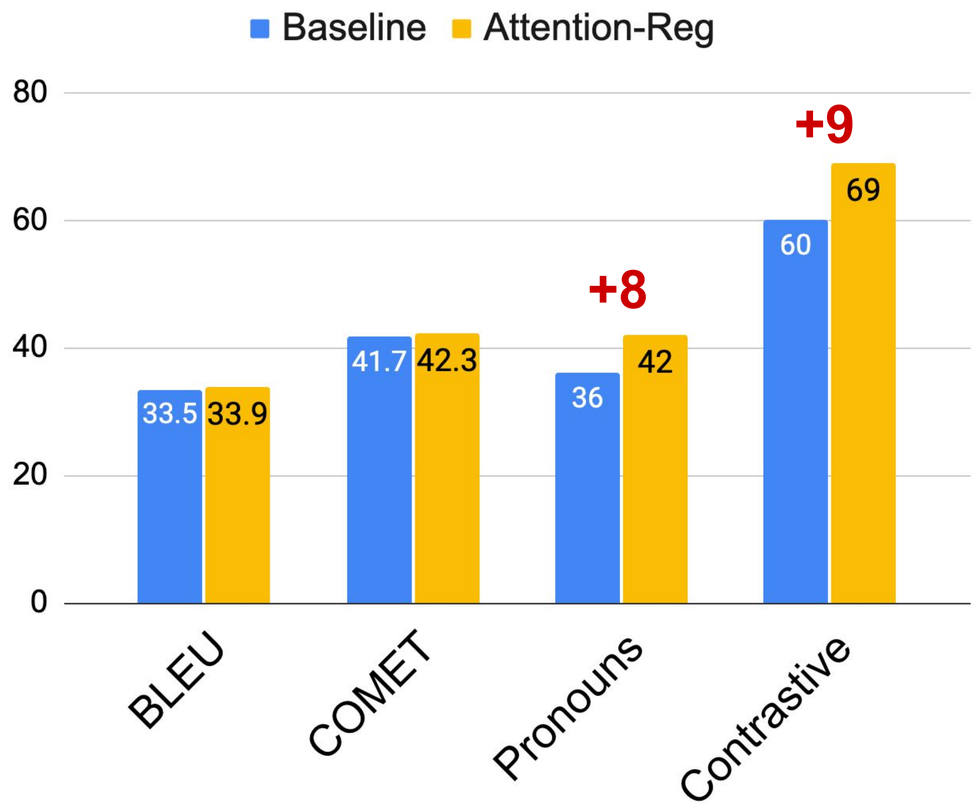

Baseline
Attention-Reg
80
60
40
20
0
33.5
41.7
36
60
BLEU
COMET
Pronouns
Contrastive

Results



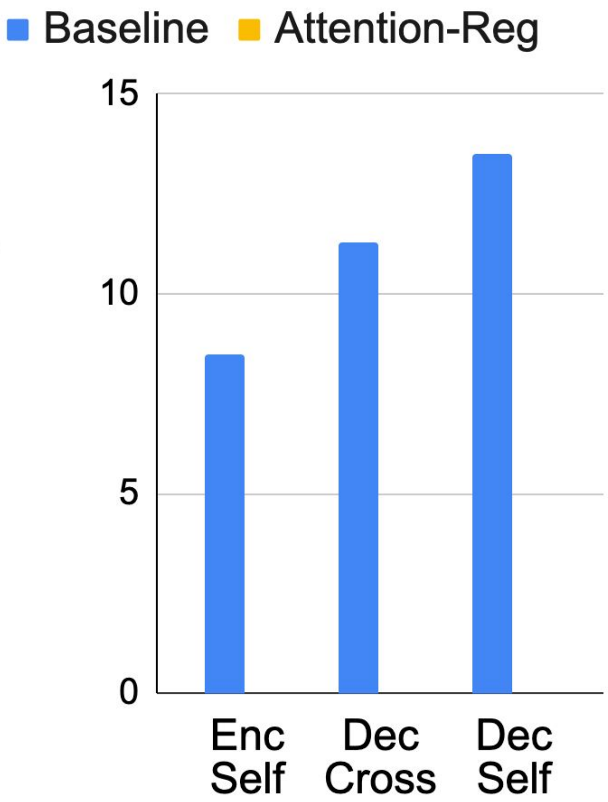
Baseline
Attention-Reg
80
60
40
20
0
33.5
33.9
41.7
42.3
36
42
60
69
BLEU
COMET
Pronouns
Contrastive

Results

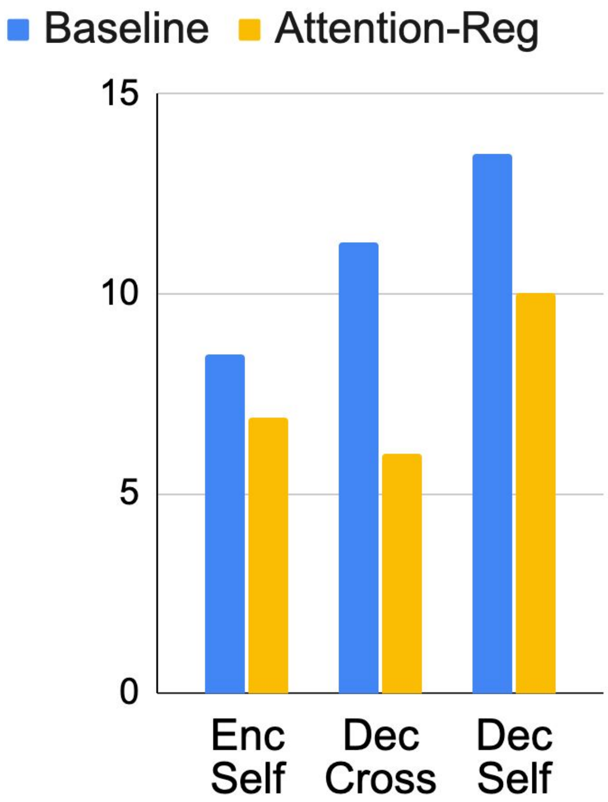


Baseline
Attention-Reg
80
60
40
20
0
+9
69
60
+8
41.7
42.3
42
36
33.5
33.9
BLEU
COMET
Pronouns
Contrastive

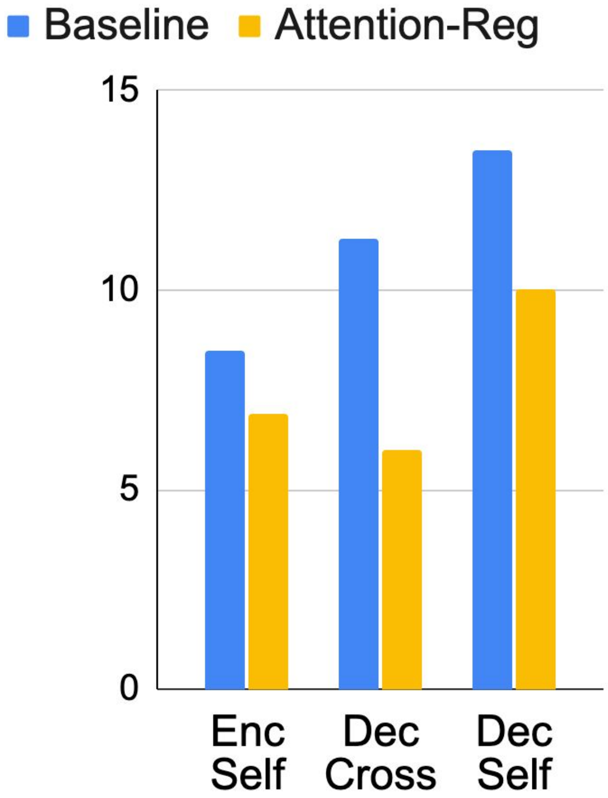
Results



Results



Results



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

Baseline

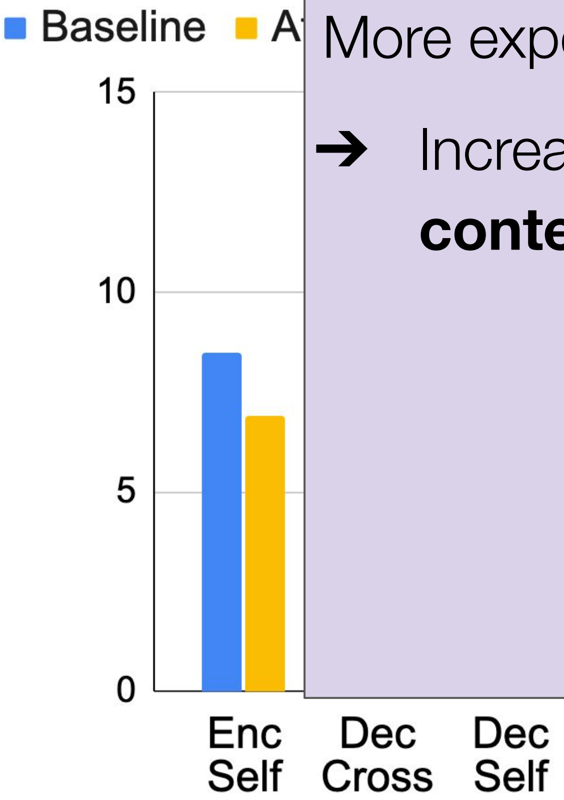
On dispose de son rapport?
Oui, **elle** est déjà à l'infirmerie.

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**

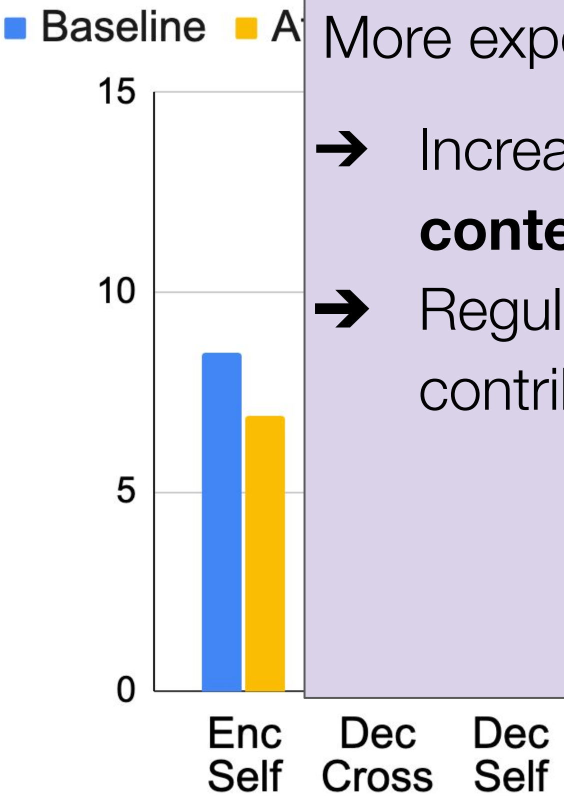
Have we got her report?
mary already.

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.

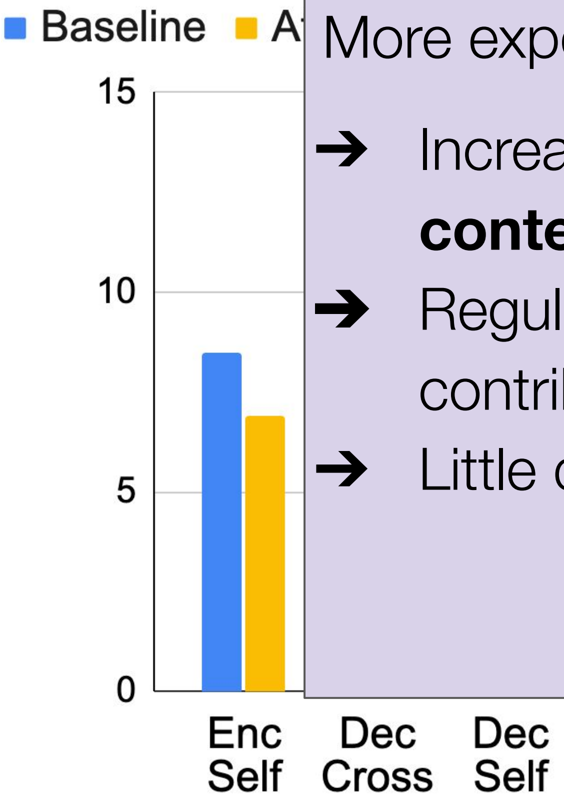
ot her report?
rmary already.

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

Results

Have we got her report?

mary already.



More experiments & results in paper:

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

son rapport?

à l'infirmerie.

got her report?

rmary already.

de son rapport?

Oui, il est déjà à l'infirmerie.

When Does Translation Require Context? A Data-driven, Multilingual Exploration

Kayo Yin*, Patrick Fernandes*, André Martins, Graham Neubig
(Ongoing work)

*Equal contribution

Evaluating Document-Level Machine Translation

- ➔ In machine translation (MT), context is crucial to translate certain discourse phenomena

Evaluating Document-Level Machine Translation

- In machine translation (MT), context is crucial to translate certain discourse phenomena

- However these phenomena represent only a small portion of the words in natural language data

Evaluating Document-Level Machine Translation

- ➔ In machine translation (MT), context is crucial to translate certain discourse phenomena
- ➔ However these phenomena represent only a small portion of the words in natural language data
- ➔ Common translation metrics don't provide a clear picture of performance in these

Evaluating Document-Level Machine Translation

➔ Recent work on context-aware MT side-steps this by using *contrastive* datasets

Evaluating Document-Level Machine Translation

- ➔ Recent work on context-aware MT side-steps this by using *contrastive* datasets
- ➔ However the availability of these datasets is limited

Evaluating Document-Level Machine Translation

- Recent work on context-aware MT side-steps this by using *contrastive* datasets
- However the availability of these datasets is limited
- Also this type of evaluation does not measure translation performance directly

Evaluating Document-Level Machine Translation

➔ In this work, we propose *data-driven*, *semi-automatic methodology* for identifying salient phenomena

Evaluating Document-Level Machine Translation

- ➔ In this work, we propose *data-driven*, *semi-automatic methodology* for identifying salient phenomena
- ➔ We create a first-of-its-kind multilingual benchmark testing these discourse phenomena

Evaluating Document-Level Machine Translation

- ➔ In this work, we propose *data-driven*, *semi-automatic methodology* for identifying salient phenomena

- ➔ We create a first-of-its-kind multilingual benchmark testing these discourse phenomena

- ➔ We evaluate multiple CAMT models, both trained by us and commercially available, on this benchmark

Measuring Context Usage

➔ Previously, we proposed *conditional cross-mutual information* (CXMI)

$$\mathrm{CXMI}(C \rightarrow Y || X) = \mathrm{H}_{q_{MT_A}}(Y||X) - \mathrm{H}_{q_{MT_C}}(Y||X, C)$$

Measuring Context Usage

- Previously, we proposed *conditional cross-mutual information* (CXMI)

$$\mathrm{CXMI}(C \rightarrow Y||X) = \mathrm{H}_{q_{MT_A}}(Y||X) - \mathrm{H}_{q_{MT_C}}(Y||X, C)$$

- This is *corpus-level* metric that tells us how well the context helps modelling a dataset

Measuring Context Usage

➔ We propose a *sentence-level* extension, Pointwise Cross Mutual Information (P-CXMI)

$$\text{P-CXMI}(y, x, C) = -\log \frac{q_{MT_A}(y|x)}{q_{MT_C}(y|x, C)}$$

Measuring Context Usage

➔ We propose a *sentence-level* extension, Pointwise Cross Mutual Information (P-CXMI)

$$\text{P-CXMI}(y, x, C) = -\log \frac{q_{MT_A}(y|x)}{q_{MT_C}(y|x, C)}$$

➔ It can also be extended to *word-level*

$$\text{P-CXMI}(i, y, x, C) = -\log \frac{q_{MT_A}(y_i|y_{t<i}, x)}{q_{MT_C}(y_i|y_{t<i}, x, C)}$$

Which Translation Phenomena Benefit from Context?

Which Translation Phenomena Benefit from Context?

➔ Look at POS tags with high mean P-CXMI

Which Translation Phenomena Benefit from Context?

➔ Look at POS tags with high mean P-CXMI

➔ Look at vocabulary items with high mean P-CXMI

Which Translation Phenomena Benefit from Context?

➔ Look at POS tags with high mean P-CXMI

➔ Look at vocabulary items with high mean P-CXMI

➔ Look at individual tokens with high P-CXMI

Which Translation Phenomena Benefit from Context?

- ~120k parallel sentences from TED talk transcripts

Which Translation Phenomena Benefit from Context?

- ~120k parallel sentences from TED talk transcripts

- 14 language pairs: English → Arabic, German, Spanish, French, Hebrew, Italian, Japanese, Korean,

  Dutch, Portuguese, Romanian, Russian, Turkish and Mandarin Chinese

Which Translation Phenomena Benefit from Context?

	ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
CXMI	0.073	0.008	0.011	0.011	0.021	0.015	0.067	0.035	0.005	0.009	0.051	0.015	0.016	0.081
P-CXMI	0.075	0.005	0.011	0.021	0.023	0.016	0.059	0.038	0.002	0.013	0.049	0.015	0.014	0.057

Which Translation Phenomena Benefit from Context?

	ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
CXMI	0.073	0.008	0.011	0.011	0.021	0.015	0.067	0.035	0.005	0.009	0.051	0.015	0.016	0.081
P-CXMI	0.075	0.005	0.011	0.021	0.023	0.016	0.059	0.038	0.002	0.013	0.049	0.015	0.014	0.057
PROPN	-0.001	-0.011	0.022	0.003	0.013	0.005	0.054	0.013	-0.006	0.003	0.114	-0.009	0.015	0.028

Which Translation Phenomena Benefit from Context?

Avelile's mother had HIV virus. Avelile had the virus, she was born with the virus. 阿维利尔的母亲是携有艾滋病病毒。 阿维利尔也有艾滋病病毒。她一生下来就有。	Lexical Cohesion

Which Translation Phenomena Benefit from Context?

	ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
CXMI	0.073	0.008	0.011	0.011	0.021	0.015	0.067	0.035	0.005	0.009	0.051	0.015	0.016	0.081
P-CXMI	0.075	0.005	0.011	0.021	0.023	0.016	0.059	0.038	0.002	0.013	0.049	0.015	0.014	0.057
PROPN	-0.001	-0.011	0.022	0.003	0.013	0.005	0.054	0.013	-0.006	0.003	0.114	-0.009	0.015	0.028
PRON.2	0.036	$\boxed{0.39}$	$\boxed{0.016}$	$\boxed{0.041}$	$\boxed{0.144}$	$\boxed{0.033}$			-0.008	$\boxed{0.318}$	0.018	$\boxed{0.109}$	-0.032	$\boxed{0.074}$

Which Translation Phenomena Benefit from Context?

Avelile's mother had HIV virus. Avelile had the virus, she was born with the virus. 阿维利尔的母亲是携有艾滋病病毒。阿维利尔也有艾滋病病毒。她一生下来就有。	Lexical Cohesion
Your daughter? Your niece? *Votre fille ?* Votre nièce ?	Formality (T-V)

Which Translation Phenomena Benefit from Context?

Avelile's mother had HIV virus. Avelile had the virus, she was born with the virus. 阿维利尔的母亲是携有艾滋病病毒。阿维利尔也有艾滋病病毒。她一生下来就有。	Lexical Cohesion
Your daughter? Your niece? *Votre fille ?* Votre nièce ?	Formality (T-V)
Roger. I got'em. Two-Six, this is Two-Six , we're mobile. 了解 捕捉した。 2-6 こちら移動中だ。	Formality (Honorifics)

Which Translation Phenomena Benefit from Context?

	ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
CXMI	0.073	0.008	0.011	0.011	0.021	0.015	0.067	0.035	0.005	0.009	0.051	0.015	0.016	0.081
P-CXMI	0.075	0.005	0.011	0.021	0.023	0.016	0.059	0.038	0.002	0.013	0.049	0.015	0.014	0.057
PROPN	-0.001	-0.011	0.022	0.003	0.013	0.005	0.054	0.013	-0.006	0.003	0.114	-0.009	0.015	0.028
PRON.2	0.036	0.39	0.016	0.041	0.144	0.033			-0.008	0.318	0.018	0.109	-0.032	0.074
VERB	0.055	0.013	0.028	0.012	0.029	0.022	0.042	0.093	0.013	0.028	0.092	0.046	0.05	0.049
PRON	0.029	0.016	0.003	0.011	0.052	0.015	0.012	0.062	0.0	0.044	0.027	0.031	0.0	0.064
PRON.1	0.019	0.021	0.01	0.029	0.034	0.025			-0.002	0.071	0.041	0.04	0.007	0.062
PRON.1.Plur	0.015	-0.002	0.025	0.01	0.106	0.0				0.079	0.015	0.042	0.047	0.067
PRON.1.Sing	0.039	0.037	0.001	0.047	-0.019	0.049				0.068	0.062	0.038	-0.02	
PRON.3	0.031	0.024	-0.004	-0.0	0.053	0.009			0.003	0.058	0.024	0.047	0.002	0.097
PRON.3.Dual	0.139													
PRON.3.Plur	0.044	0.023	0.001	-0.015	0.065	0.075				0.091	0.048	0.031	0.019	0.1
PRON.3.Sing	0.026	0.024	0.008	0.008	0.056	0.006				0.037	0.034	0.059	-0.002	

Which Translation Phenomena Benefit from Context?

Avelile's mother had HIV virus. Avelile had the virus, she was born with the virus. 阿维利尔的母亲是携有艾滋病病毒。阿维利尔也有艾滋病病毒。她一生下来就有。	Lexical Cohesion
Your daughter? Your niece? *Votre fille ?* Votre nièce ?	Formality (T-V)
Roger. I got'em. Two-Six, this is Two-Six , we're mobile. 了解 捕捉した。 2-6 こちら移動中だ。	Formality (Honorifics)
Our tools today don't look like shovels and picks. They look like the stuff we walk around with. *As ferramentas de hoje não se parecem com pás e picaretas.* Elas se parecem com as coisas que usamos.	Pronouns

Which Translation Phenomena Benefit from Context?

	ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
CXMI	0.073	0.008	0.011	0.011	0.021	0.015	0.067	0.035	0.005	0.009	0.051	0.015	0.016	0.081
P-CXMI	0.075	0.005	0.011	0.021	0.023	0.016	0.059	0.038	0.002	0.013	0.049	0.015	0.014	0.057
PROPN	-0.001	-0.011	0.022	0.003	0.013	0.005	0.054	0.013	-0.006	0.003	0.114	-0.009	0.015	0.028
PRON.2	0.036	0.39	0.016	0.041	0.144	0.033			-0.008	0.318	0.018	0.109	-0.032	0.074
VERB	0.055	0.013	0.028	0.012	0.029	0.022	0.042	0.093	0.013	0.028	0.092	0.046	0.05	0.049
PRON	0.029	0.016	0.003	0.011	0.052	0.015	0.012	0.062	0.0	0.044	0.027	0.031	0.0	0.064
PRON.1	0.019	0.021	0.01	0.029	0.034	0.025			-0.002	0.071	0.041	0.04	0.007	0.062
PRON.1.Plur	0.015	-0.002	0.025	0.01	0.106	0.0				0.079	0.015	0.042	0.047	0.067
PRON.1.Sing	0.039	0.037	0.001	0.047	-0.019	0.049				0.068	0.062	0.038	-0.02	
PRON.2	0.036	0.39	0.016	0.041	0.144	0.033			-0.008	0.318	0.018	0.109	-0.032	0.074
PRON.2.Plur	0.05	-1.203	-0.062	0.017	0.095	0.014					0.022	0.051	-0.033	
PRON.2.Sing	0.02	0.412	0.061	0.406	0.226	0.089				0.318	0.007	0.662	-0.027	
PRON.3	0.031	0.024	-0.004	-0.0	0.053	0.009			0.003	0.058	0.024	0.047	0.002	0.097
PRON.3.Dual	0.139													
PRON.3.Plur	0.044	0.023	0.001	-0.015	0.065	0.075				0.091	0.048	0.031	0.019	0.1
PRON.3.Sing	0.026	0.024	0.008	0.008	0.056	0.006				0.037	0.034	0.059	-0.002	
VERB.Fut			-0.007	-0.069	0.009	0.061				0.044		0.012	0.034	
VERB.Imp			0.102	0.024		0.044				0.118	0.18			
VERB.Past		0.075	0.032	0.019	0.053	0.041			0.064	0.046	0.029	0.115	0.047	
VERB.Pres		0.017	0.029	0.014		0.022			0.002	0.024	0.083	0.022	0.051	

Which Translation Phenomena Benefit from Context?

Avelile's mother had HIV virus. Avelile had the virus, she was born with the virus. 阿维利尔的母亲是携有艾滋病病毒。 阿维利尔也有艾滋病病毒。她一生下来就有。	Lexical Cohesion
Your daughter? Your niece? *Votre fille ?* Votre nièce ?	Formality (T-V)
Roger. I got'em. Two-Six, this is Two-Six , we're mobile. 了解 捕捉した。 2-6 こちら移動中だ。	Formality (Honorifics)
Our tools today don't look like shovels and picks. They look like the stuff we walk around with. *As ferramentas de hoje não se parecem com pás e picaretas.* Elas se parecem com as coisas que usamos.	Pronouns
Louis XIV had a lot of people working for him. They made his silly outfits, like this. *Luis XIV tenía un montón de gente trabajando para él.* Ellos hacían sus trajes tontos, como éste.	Verb Form

Which Translation Phenomena Benefit from Context?

Avelile's mother had HIV virus. Avelile had the virus, she was born with the virus. 阿维利尔的母亲是携有艾滋病病毒。 阿维利尔也有艾滋病病毒。她一生下来就有。	Lexical Cohesion
Your daughter? Your niece? *Votre fille ?* Votre nièce ?	Formality (T-V)
Roger. I got'em. Two-Six, this is Two-Six , we're mobile. 了解 捕捉した。 2-6 こちら移動中だ。	Formality (Honorifics)
Our tools today don't look like shovels and picks. They look like the stuff we walk around with. *As ferramentas de hoje não se parecem com pás e picaretas.* Elas se parecem com as coisas que usamos.	Pronouns
Louis XIV had a lot of people working for him. They made his silly outfits, like this. *Luis XIV tenía un montón de gente trabajando para él.* Ellos hacían sus trajes tontos, como éste.	Verb Form
They're the ones who know what society is going to be like in another generation. I don't. *Ancak onlar başka bir nesilde toplumun nasıl olacağını biliyorlar.* Ben bilmiyorum.	Ellipsis

Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

➔ **Lexical Cohesion:** tag target words *y* if the aligned source and target words pair *(x,y)* appears at least 3 times in the document

Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

- **Lexical Cohesion:** tag target words *y* if the aligned source and target words pair *(x,y)* appears at least 3 times in the document
- **Formality:** tag target words that are T-V pronouns/verbs or honorific terms

Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

- ➔ **Lexical Cohesion:** tag target words *y* if the aligned source and target words pair *(x,y)* appears at least 3 times in the document
- ➔ **Formality:** tag target words that are T-V pronouns/verbs or honorific terms
- ➔ **Pronoun choice:** tag target pronouns if the corresponding source pronoun has multiple possible translations

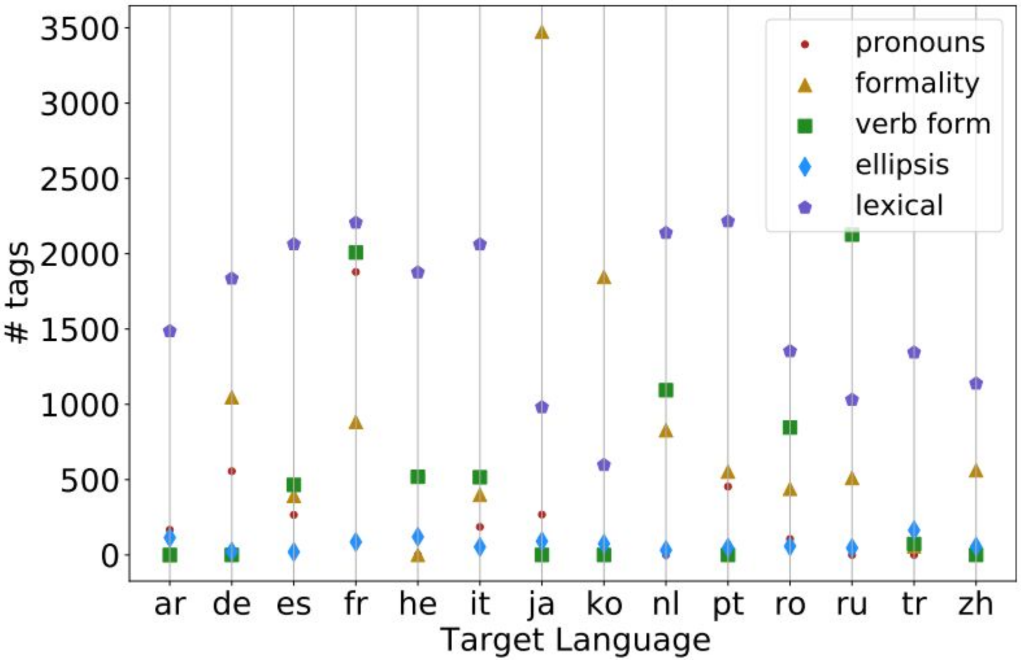
Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

- ➔ **Lexical Cohesion:** tag target words *y* if the aligned source and target words pair *(x,y)* appears at least 3 times in the document
- ➔ **Formality:** tag target words that are T-V pronouns/verbs or honorific terms
- ➔ **Pronoun choice:** tag target pronouns if the corresponding source pronoun has multiple possible translations
- ➔ **Verb form**: tag target verbs if it has a verb form such that the corresponding source verb form has multiple possible translations

Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

- **Lexical Cohesion:** tag target words *y* if the aligned source and target words pair *(x,y)* appears at least 3 times in the document
- **Formality:** tag target words that are T-V pronouns/verbs or honorific terms
- **Pronoun choice:** tag target pronouns if the corresponding source pronoun has multiple possible translations
- **Verb form**: tag target verbs if it has a verb form such that the corresponding source verb form has multiple possible translations
- **Ellipsis:** tag target verbs, nouns and pronouns if the source sentence contains an ellipsis and the target word is not aligned to any source word

Multilingual **D**iscourse-**A**ware (MuDA) Benchmark



Multilingual **D**iscourse-**A**ware (MuDA) Benchmark

	lexical	formality	pronouns	verb form	ellipsis
de	1.00	0.74	0.70	–	0.54
es	1.00	0.92	1.00	1.00	0.53
fr	1.00	1.00	0.96	0.92	0.43
ja	1.00	0.98	1.00	–	0.41
ko	1.00	0.93	–	–	0.26
pt	0.99	0.88	1.00	–	0.31
ru	1.00	1.00	–	0.96	0.50
tr	1.00	1.00	–	1.00	0.57
zh	1.00	1.00	–	–	0.78

Table 3: Precision of MuDA tags on 50 utterances.

A Cross-lingual, Cross-Model Exploration of Context-aware MT

- We evaluate a sentence-level MT model and context-aware MT model on our system
  - We use a transformer small
  - For the context-aware method, we *prepend* the previous target context sentences to the current target

A Cross-lingual, Cross-Model Exploration of Context-aware MT

		ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
BLEU	no-context	15.69	31.02	38.16	**27.09**	25.29	34.91	4.64	8.15	**35.23**	39.83	27.6	19.7	17.12	**17.24**
	context	14.93	31.06	38.51	26.62	25.96	35.02	3.18	8.62	35.03	39.89	27.09	19.66	17.15	15.59
	context-gold	**17.15**	**31.08**	**38.57**	26.93	**26.36**	**35.25**	**5.63**	**8.87**	35.11	**40.08**	**29.84**	**19.98**	**17.4**	16.92
COMET	no-context	**0.113**	**0.152**	0.422	–0.057	**0.300**	0.312	–0.876	–0.148	0.310	**0.526**	**0.426**	0.029	**0.232**	**-0.100**
	context	0.055	0.130	**0.424**	**-0.047**	0.273	0.319	–0.914	–0.069	0.314	0.525	0.398	–0.001	0.211	–0.192
	context-gold	0.092	0.129	**0.424**	–0.049	0.276	**0.323**	**-0.810**	**-0.049**	**0.317**	0.523	0.396	**-0.001**	0.213	–0.150
all	no-context	0.512	0.65	0.694	0.63	0.627	0.635	0.287	0.37	0.678	0.688	0.592	0.529	0.462	0.402
	context	0.501	0.65	0.695	0.64	0.63	0.637	0.209	0.379	0.678	0.688	0.589	0.528	0.464	0.364
	context-gold	0.524	0.65	0.695	0.641	0.631	0.639	0.295	0.385	0.679	0.69	0.616	0.531	0.464	0.409
ellipsis	no-context	0.34	0.372	0.286	0.226	0.387	0.355	0.033	0.159	0.314	0.436	0.172	0.25	0.171	0.146
	context	0.318	0.278	0.303	0.209	0.392	0.339	0.026	0.195	0.273	0.421	0.239	0.145	0.132	0.09
	context-gold	0.364	0.235	0.333	0.202	0.4	0.323	0.031	0.192	0.273	0.464	0.25	0.104	0.13	0.148
formality	no-context	–	0.631	0.29	0.748	–	0.328	0.405	0.138	0.665	0.619	0.433	0.451	0.165	0.689
	context	–	0.623	0.325	0.745	–	0.362	0.369	0.135	0.669	0.607	0.428	0.476	0.204	0.693
	context-gold	–	0.649	0.317	0.74	–	0.347	0.401	0.141	0.677	0.612	0.422	0.471	0.271	0.697
lexical	no-context	0.633	0.742	0.815	0.816	0.713	0.75	0.591	0.515	0.822	0.852	0.689	0.61	0.672	0.612
	context	0.621	0.733	0.813	0.812	0.717	0.764	0.595	0.539	0.82	0.855	0.669	0.586	0.636	0.552
	context-gold	0.657	0.736	0.819	0.816	0.726	0.769	0.607	0.577	0.821	0.857	0.704	0.591	0.645	0.568
pronouns	no-context	0.57	0.574	0.575	0.718	–	0.512	0.363	–	–	0.461	0.402	–	–	–
	context	0.569	0.57	0.56	0.733	–	0.548	0.362	–	–	0.44	0.359	–	–	–
	context-gold	0.588	0.579	0.565	0.738	–	0.536	0.345	–	–	0.466	0.351	–	–	–
verb tense	no-context	–	–	0.266	0.389	0.258	0.291	–	–	0.479	–	0.289	0.213	0.128	–
	context	–	–	0.261	0.397	0.254	0.312	–	–	0.472	–	0.305	0.212	0.079	–
	context-gold	–	–	0.261	0.398	0.263	0.307	–	–	0.478	–	0.337	0.227	0.09	–

Table 4: BLEU, COMET, and Word f-meas per tag for our `base` context-aware models. Best BLEU and COMET are **bolded** whereas word f-meas higher than no-context by > 0.025 are underlined.

A Cross-lingual, Cross-Model Exploration of Context-aware MT

- To evaluate more powerful models, we also finetune a large, pretrained model on this task
  - We do this for DE, FR, JA and ZH
  - We use a transformer large
  - We pretrain on Paracrawl, JParacrawl and Backtranslated News

A Cross-lingual, Cross-Model Exploration of Context-aware MT

		de	fr	ja	zh
BLEU	no-context	37.7	50.23	16.39	23.07
	context	38.23	50.47	12.87	23.32
	context-gold	**38.77**	**51.64**	**17.44**	**23.8**
COMET	no-context	0.483	0.628	0.135	0.249
	context	0.486	0.632	–0.004	0.271
	context-gold	**0.493**	**0.645**	**0.153**	**0.287**
all	no-context	0.697	0.733	0.474	0.447
	context	0.699	0.734	0.427	0.456
	context-gold	0.704	0.741	0.475	0.463
ellipsis	no-context	0.421	0.447	0.227	0.195
	context	<u>0.485</u>	0.415	0.085	0.191
	context-gold	<u>0.457</u>	0.38	0.152	0.209
formality	no-context	0.632	0.797	0.506	0.724
	context	0.654	0.792	0.495	0.736
	context-gold	<u>0.698</u>	0.811	0.527	0.719
lexical	no-context	0.774	0.865	0.682	0.648
	context	0.776	0.862	0.677	0.626
	context-gold	0.795	0.872	<u>0.73</u>	0.644
pronouns	no-context	0.623	0.755	0.485	–
	context	0.613	0.76	0.481	–
	context-gold	0.645	0.778	0.492	–
verb tense	no-context	–	0.518	–	–
	context	–	0.517	–	–
	context-gold	–	0.53	–	–

Table 5: Word f-meas per tag for our `large` models. Best BLEU and COMET are **bolded** whereas word f-meas higher than no-context by > 0.025 are <u>underlined</u>.

A Cross-lingual, Cross-Model Exploration of Context-aware MT

- Finally we consider two commercial engines and evaluate them on our benchmark
    - the *Google Cloud Translation v2* API
    - the *DeepL v2 API*

A Cross-lingual, Cross-Model Exploration of Context-aware MT

		ar	de	es	fr	he	it	ja	ko	nl	pt	ro	ru	tr	zh
BLEU	Google	19.45	36.52	41.29	35.09	29.37	35.60	11.76	11.22	38.51	**45.99**	26.61	19.61	21.30	24.22
	DeepL (sent)	x	38.49	40.43	43.48	x	37.07	13.85	x	40.92	39.41	32.67	25.98	x	28.10
	Deepl (doc)	x	**39.21**	**42.75**	**45.09**	x	**40.54**	13.86	x	**41.11**	40.64	**33.24**	**29.08**	x	**28.93**
COMET	Google	0.464	0.448	0.722	0.567	0.554	37.070	**0.208**	0.405	0.594	**0.775**	0.682	0.491	0.663	0.299
	DeepL (sent)	x	**0.498**	0.734	0.628	x	0.658	0.138	x	0.589	0.734	0.778	0.510	x	0.352
	Deepl (doc)	x	0.474	**0.747**	**0.653**	x	**0.671**	0.206	x	**0.602**	0.602	**0.790**	**0.529**	x	**0.362**
all	Google	0.563	0.69	0.748	0.72	0.652	0.676	0.412	0.422	0.683	0.667	0.573	0.491	0.531	0.445
	DeepL (sent)	x	0.705	0.737	0.732	x	0.676	0.454	x	0.706	0.652	0.638	0.602	x	0.528
	DeepL (doc)	x	0.706	0.742	0.74	x	0.681	0.451	x	0.711	0.668	0.642	0.611	x	0.542
ellipsis	Google	0.376	0.462	0.414	0.453	0.481	0.377	0.209	0.254	0.381	0.549	0.314	0.333	0.271	0.193
	DeepL (sent)	x	0.462	0.444	0.482	x	0.467	0.299	x	0.439	0.407	0.36	0.312	x	0.265
	DeepL (doc)	x	0.462	0.5	0.537	x	0.483	0.291	x	0.381	0.407	0.372	0.279	x	0.261
formality	Google	x	0.579	0.266	0.727	x	0.279	0.483	0.099	0.624	0.633	0.449	0.488	0.326	0.29
	DeepL (sent)	x	0.665	0.281	0.655	x	0.332	0.419	x	0.622	0.584	0.521	0.522	x	0.722
	DeepL (doc)	x	0.66	0.272	0.765	x	0.35	0.455	x	0.631	0.58	0.52	0.549	x	0.729
lexical	Google	0.663	0.767	0.856	0.852	0.711	0.789	0.568	0.597	0.82	0.856	0.686	0.592	0.662	0.698
	DeepL (sent)	x	0.77	0.822	0.851	x	0.777	0.628	x	0.807	0.842	0.713	0.619	x	0.679
	DeepL (doc)	x	0.782	0.839	0.865	x	0.779	0.629	x	0.801	0.846	0.721	0.637	x	0.673
pronouns	Google	0.64	0.622	0.618	0.741	–	0.509	0.467	–	–	0.503	0.436	–	–	–
	DeepL (sent)	x	0.62	0.554	0.707	x	0.509	0.5	x	–	0.47	0.473	–	x	–
	DeepL (doc)	x	0.66	0.571	0.75	x	0.517	0.555	x	–	0.497	0.502	–	x	–
verb tense	Google	–	–	0.399	0.524	0.265	0.41	–	–	0.515	–	0.345	0.312	0.204	–
	DeepL (sent)	x	–	0.415	0.548	x	0.455	–	x	0.547	–	0.409	0.328	x	–
	DeepL (doc)	x	–	0.432	0.549	x	0.46	–	x	0.568	–	0.409	0.346	x	–

Table 6: Scores for commercial models. Best BLEU and COMET are **bolded**, DeepL (doc) where word f-meas is higher than DeepL (sent) by >0.025 are underlined. Languages not supported are ‘x’ed.

Signed Coreference Resolution

Kayo Yin, Kenneth DeHaan, Malihe Alikhani

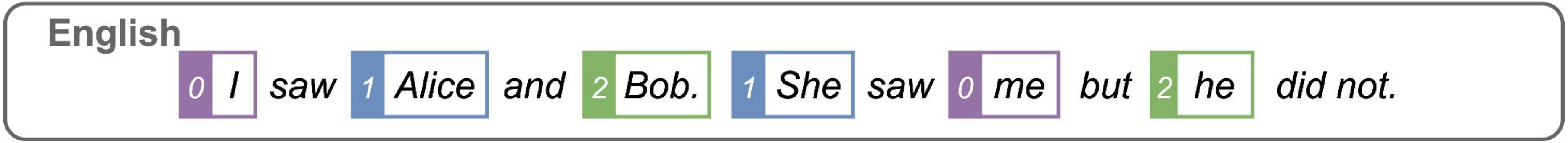
(EMNLP 2021)

Coreference Resolution

English

I saw Alice and Bob. She saw me but he did not.

Coreference Resolution



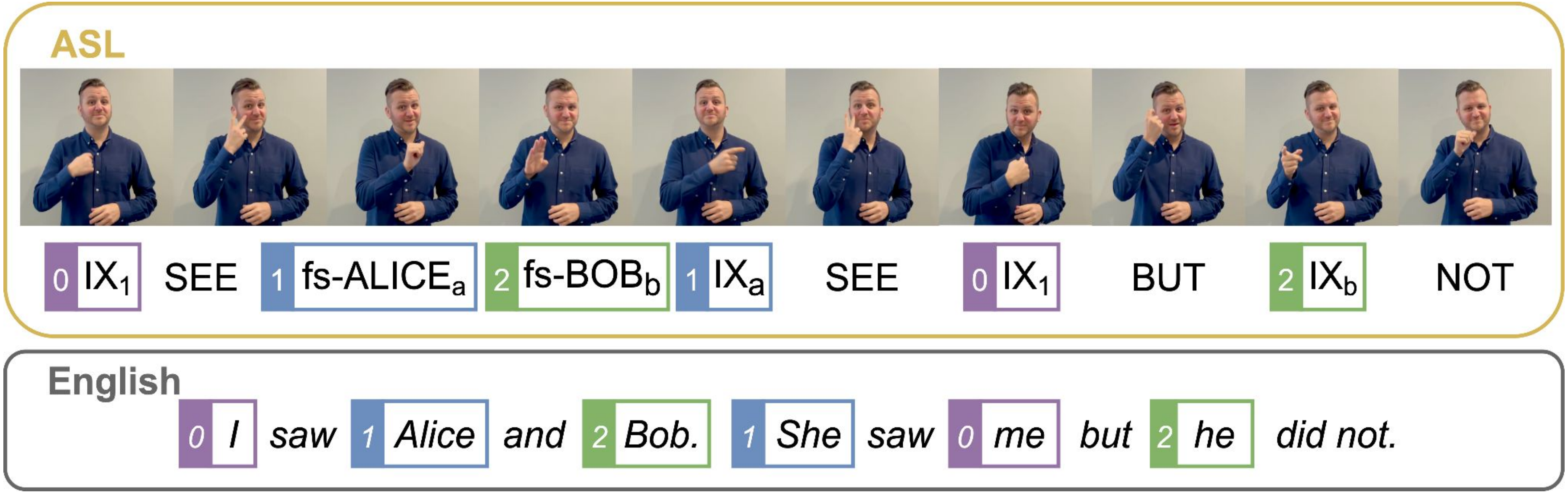
Signed Coreference Resolution

ASL

English

0 I saw 1 Alice and 2 Bob. 1 She saw 0 me but 2 he did not.

Signed Coreference Resolution



Signed Coreference Resolution

➔ Novel challenges in modeling **discourse** and **spatial context**

Signed Coreference Resolution

➔ Novel challenges in modeling **discourse** and **spatial context**

➔ Better understanding of **grounding** in different forms of communication

Signed Coreference Resolution

➔ Novel challenges in modeling **discourse** and **spatial context**

➔ Better understanding of **grounding** in different forms of communication

➔ Broaden the scope of NLP to **multiple modalities**

Signed Coreference Resolution

- ➔ Novel challenges in modeling **discourse** and **spatial context**
- ➔ Better understanding of **grounding** in different forms of communication
- ➔ Broaden the scope of NLP to **multiple modalities**
- ➔ Enable **Sign Language Processing** technologies

Outline

1. Pronominal Pointing Signs
2. Signed Coreference Resolution
3. Unsupervised Continuous Multigraph
4. Results & Discussion

Outline

1. Pronominal Pointing Signs
2. Signed Coreference Resolution
3. Unsupervised Continuous Multigraph
4. Results & Discussion

Pronominal Pointing Signs

➔ Pointing signs with a **pronominal** function

Pronominal Pointing Signs

➔ Pointing signs with a **pronominal** function

➔ Referents are established in the **signing space**

Pronominal Pointing Signs

➔ Pointing signs with a **pronominal** function

➔ Referents are established in the **signing space**

➔ Point to the **actual location** of the referent

Pronominal Pointing Signs

➔ Pointing signs with a **pronominal** function

➔ Referents are established in the **signing space**

➔ Point to the **actual location** of the referent

➔ Assign a **locus** to the referent

Pronominal Pointing Signs

ASL

0 IX_1 SEE 1 fs-$ALICE_a$ 2 fs-BOB_b 1 IX_a SEE 0 IX_1 BUT 2 IX_b NOT

English

0 *I* *saw* 1 *Alice* *and* 2 *Bob.* 1 *She* *saw* 0 *me* *but* 2 *he* *did not.*

Pronominal Pointing Signs

ASL

0 IX_1 SEE 1 fs-$ALICE_a$ 2 fs-BOB_b 1 IX_a SEE 0 IX_1 BUT 2 IX_b NOT

English

0 *I* *saw* 1 *Alice* *and* 2 *Bob.* 1 *She* *saw* 0 *me* *but* 2 *he* *did not.*

Complexities of Pointing Signs

➔ Pointing signs can serve **other** functions

Complexities of Pointing Signs

➔ Pointing signs can serve **other** functions

➔ Difficult to distinguish between different pointing signs based solely on **local visual features**

Complexities of Pointing Signs

English Pronouns

+ Carry some meaning on its own

ASL Pointing Signs

Complexities of Pointing Signs

<u>English Pronouns</u>

\+ Carry some meaning on its own

<u>ASL Pointing Signs</u>

\- Use the same handshape, harder to distinguish on its own

Complexities of Pointing Signs

English Pronouns

+ Carry some meaning on its own
- The same word can refer to multiple entities at once

ASL Pointing Signs

- Use the same handshape, harder to distinguish on its own

Complexities of Pointing Signs

English Pronouns

+ Carry some meaning on its own
- The same word can refer to multiple entities at once

ASL Pointing Signs

- Use the same handshape, harder to distinguish on its own
+ 1 locus = 1 referent

Complexities of Pointing Signs

English Pronouns

\+ Carry some meaning on its own
\- The same word can refer to multiple entities at once

ASL Pointing Signs

\- Use the same handshape, harder to distinguish on its own
\+ 1 locus = 1 referent
\- Loci can be reassigned to different referents

Complexities of Pointing Signs

English Pronouns

+ Carry some meaning on its own
- The same word can refer to multiple entities at once

ASL Pointing Signs

- Use the same handshape, harder to distinguish on its own
+ 1 locus = 1 referent
- Loci can be reassigned to different referents
- Referents can be assigned multiple loci

Why study Signed Coreference Resolution in NLP?

➔ Theories of coreference in spoken languages may be **extended** to signed languages

Why study Signed Coreference Resolution in NLP?

- ➔ Theories of coreference in spoken languages may be **extended** to signed languages
  - ◆ Discourse Representation Theory (Kamp et al., 2011; Steinbach and Onea 2016)

Why study Signed Coreference Resolution in NLP?

➔ Theories of coreference in spoken languages may be **extended** to signed languages

- Discourse Representation Theory (Kamp et al., 2011; Steinbach and Onea 2016)
- First mention effect (Gernsbacher and Hargreaves, 1988; Wienholz et al., 2020)

Why study Signed Coreference Resolution in NLP?

- ➔ Theories of coreference in spoken languages may be **extended** to signed languages
  - ◆ Discourse Representation Theory (Kamp et al., 2011; Steinbach and Onea 2016)
  - ◆ First mention effect (Gernsbacher and Hargreaves, 1988; Wienholz et al., 2020)
- ➔ It can help us better understand **multimodal** communication

Why study Signed Coreference Resolution in NLP?

- ➔ Theories of coreference in spoken languages may be **extended** to signed languages
  - ◆ Discourse Representation Theory (Kamp et al., 2011; Steinbach and Onea 2016)
  - ◆ First mention effect (Gernsbacher and Hargreaves, 1988; Wienholz et al., 2020)
- ➔ It can help us better understand **multimodal** communication
  - ◆ Spatial iconicity and situated referents in signed languages

Why study Signed Coreference Resolution in NLP?

- ➔ Theories of coreference in spoken languages may be **extended** to signed languages
  - ◆ Discourse Representation Theory (Kamp et al., 2011; Steinbach and Onea 2016)
  - ◆ First mention effect (Gernsbacher and Hargreaves, 1988; Wienholz et al., 2020)
- ➔ It can help us better understand **multimodal** communication
  - ◆ Spatial iconicity and situated referents in signed languages
- ➔ Widen the **accessibility** of language technologies

Outline

1. Pronominal Pointing Signs
2. Signed Coreference Resolution
3. Unsupervised Continuous Multigraph
4. Results & Discussion

Signed Coreference Resolution

Signed Coreference Resolution

IX_1 SEE fs-$ALICE_a$ fs-BOB_b IX_a SEE IX_1 BUT IX_b NOT

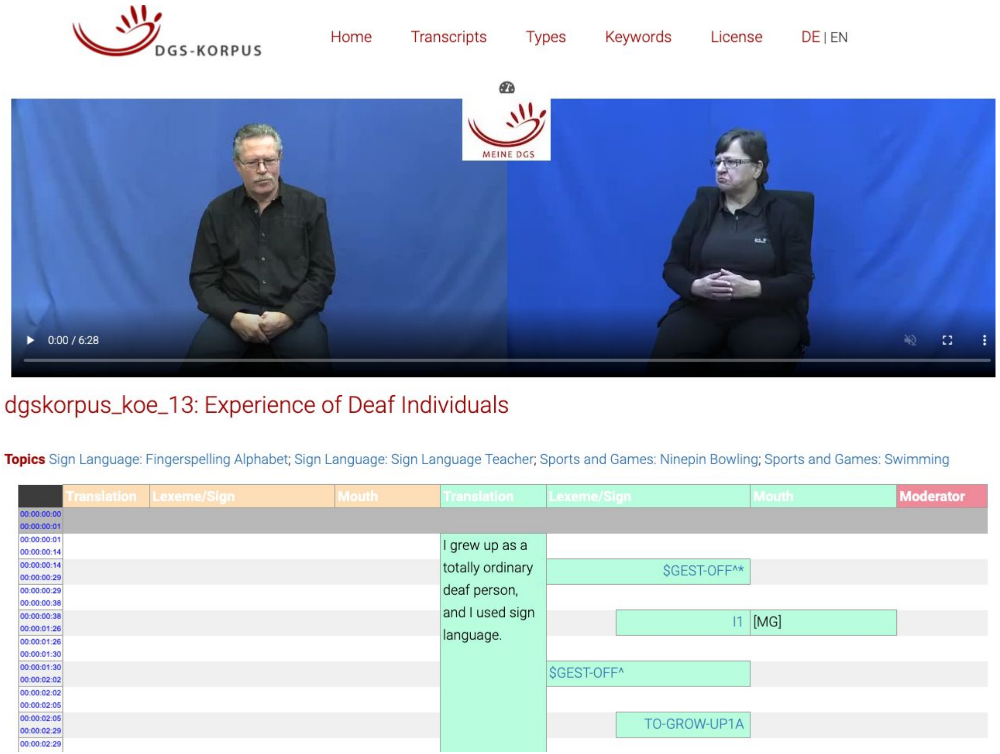
1. Mention Detection

Signed Coreference Resolution

0 IX_1 SEE 1 fs-$ALICE_a$ 2 fs-BOB_b 1 IX_a SEE 0 IX_1 BUT 2 IX_b NOT

2. Coreference Resolution

DGS-Coref Dataset



Public DGS Corpus (Hanke et al., 2020)

DGS-Coref Dataset

Task 1 (Video b'1429737', 84) - Example 61

Video: https://www.sign-lang.uni-hamburg.de/meinedgs/html/1429737_en.html#t00053952

English context:

A: Now I have knee and back pain.

A: That's why I had to stop.

A: I was active in the club for over ten years.

A: Oh well.

A: I haven't done sports actively here in North Rhine-Westphalia.

A: I’m working as a sign language teacher.

A: Back in Berlin I didn't work as a sign language teacher.

English:

A: When I came here, my partner told me that I would be a great sign language teacher.

English context you highlighted:

Reset Highlights

English sentence you highlighted:

Reset Highlights

Glosses context:

NOW1* I2 KNEE1A* PAIN3 $GEST-OFF^* LOWER-BACK1E PAIN3

I1 FINISH1

OVER-OR-ABOUT1* YEAR1A* ACTIVE1 I1

$GEST-OFF^*

HERE1 NOT1*

TO-SIGN1A LECTURER1

PAST-OR-BACK-THEN1* BERLIN1A* $INDEX1 I1 TO-SIGN1A LECTURER1 NOT3A I1*

Glosses:

$INDEX1 THROUGH2A TO-COME1 $INDEX1* $GEST-DECLINE1^ MY1* LIFE-PARTNER1 $INDEX1 TO-RECOMMEND1A* TO-SAY1 TO-MATCH1 TO-SIGN1A TO-MATCH1

Gloss context you highlighted:

- BERLIN1A*
- $INDEX1

Reset Highlights

Gloss sentence you highlighted:

Reset Highlights

How confident are you?

Not at all | Somewhat | Very

DGS-Coref Dataset

- ➔ 16m30s of signing
- ➔ 3 conversations
- ➔ 5 different signers
- ➔ 288 signed sentences
- ➔ 1,457 glosses
    - ◆ 95 <I> signs
    - ◆ 8 <YOU> signs
    - ◆ 93 <INDEX> signs

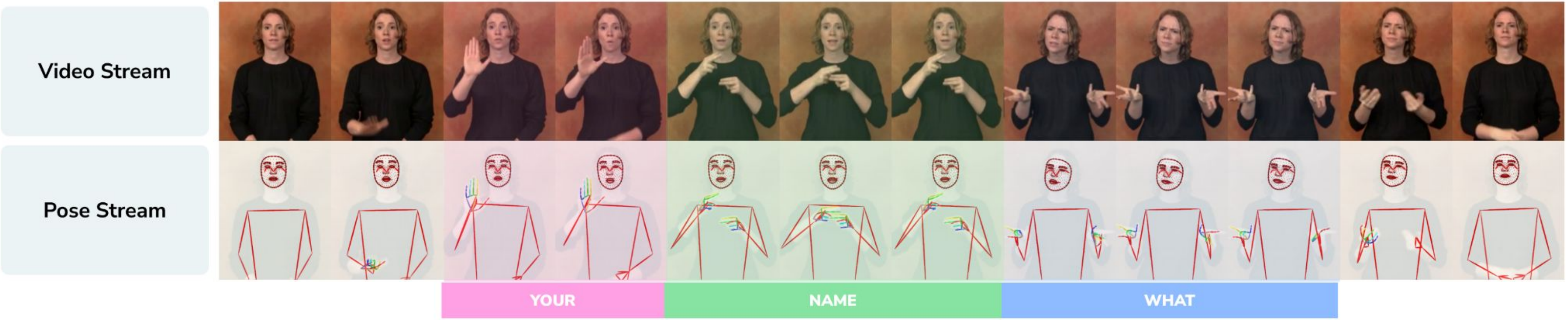
A: WITH TRIP **INDEX** SHIP **INDEX**

A: We went there with an excursion boat.

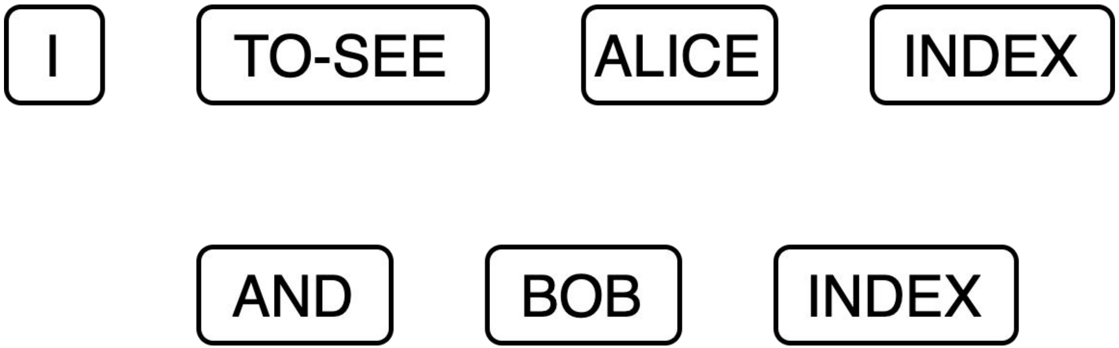
Outline

1. Pronominal Pointing Signs
2. Signed Coreference Resolution Task & Data
3. Unsupervised Continuous Multigraph
4. Results & Discussion

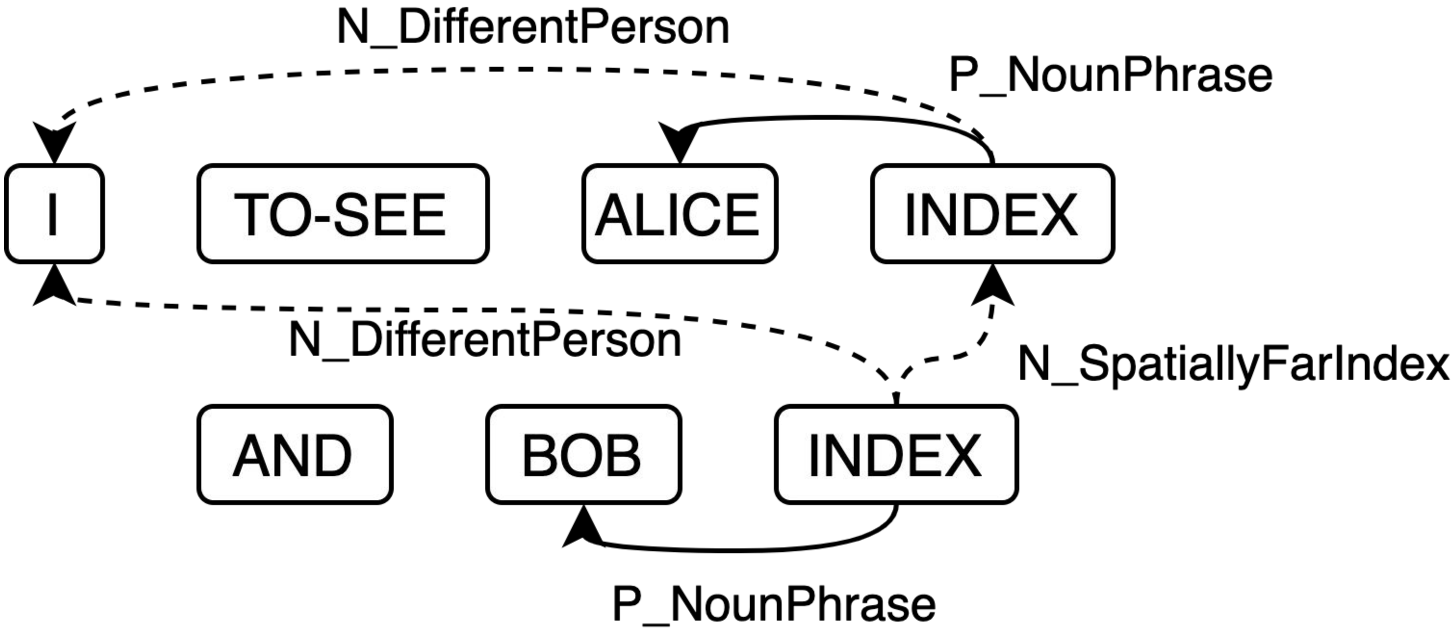
Unsupervised Continuous Multigraph



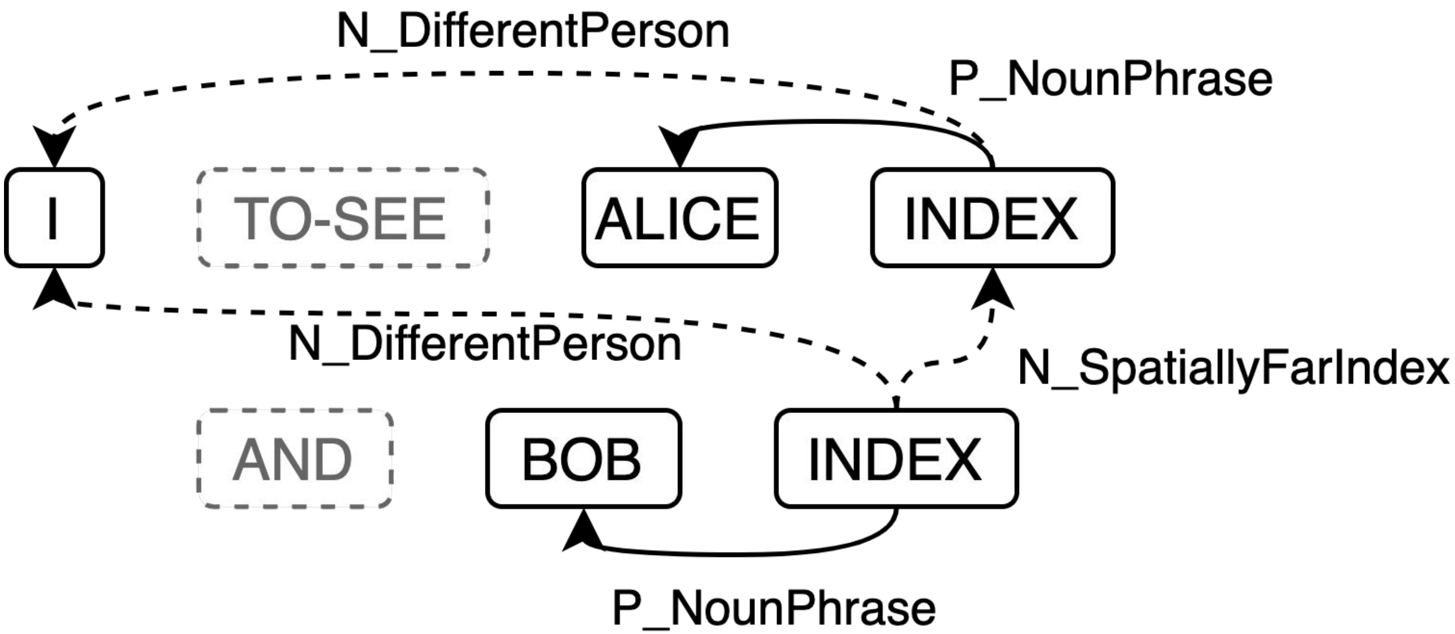
Unsupervised Continuous Multigraph



Unsupervised Continuous Multigraph



Unsupervised Continuous Multigraph

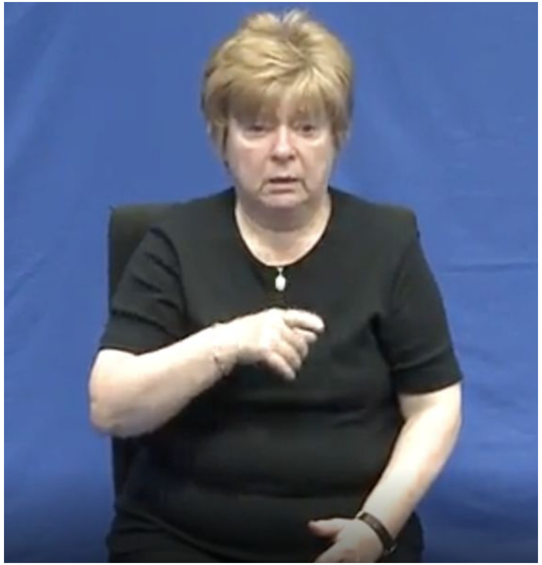


Positive Relations

1. I and I

Positive Relations

1. I and I
2. You and You

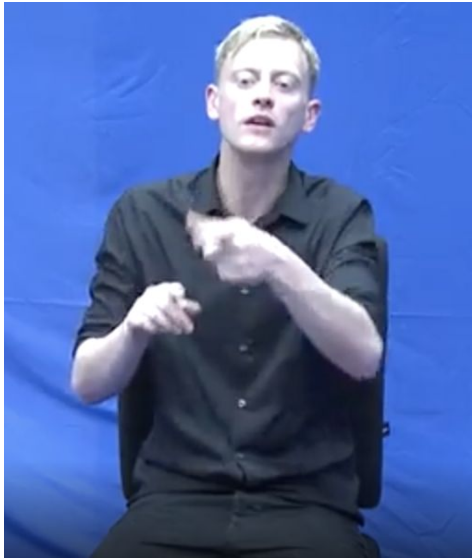


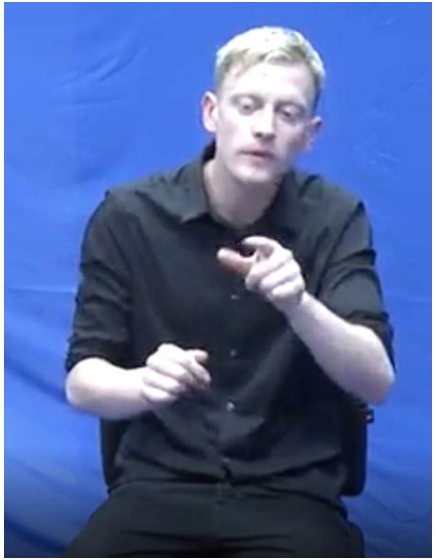
Positive Relations

1. I and I
2. You and You
3. I and You

Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index



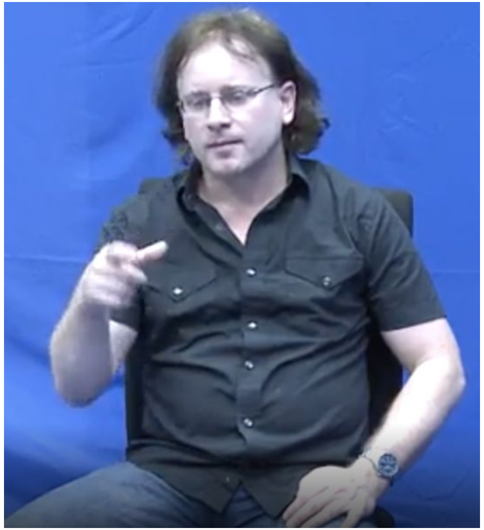


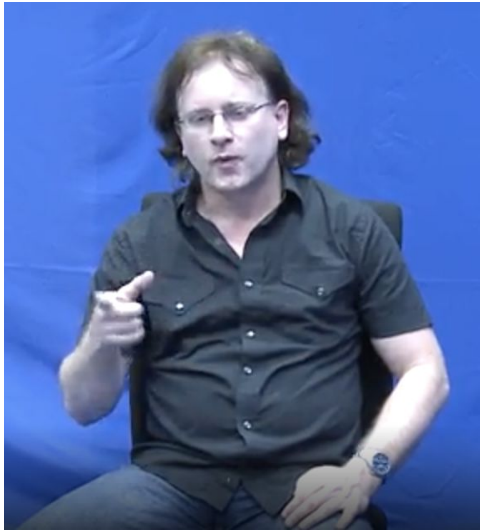
Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index
5. Noun Phrase

Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index
5. Noun Phrase
6. Spatially Close Index



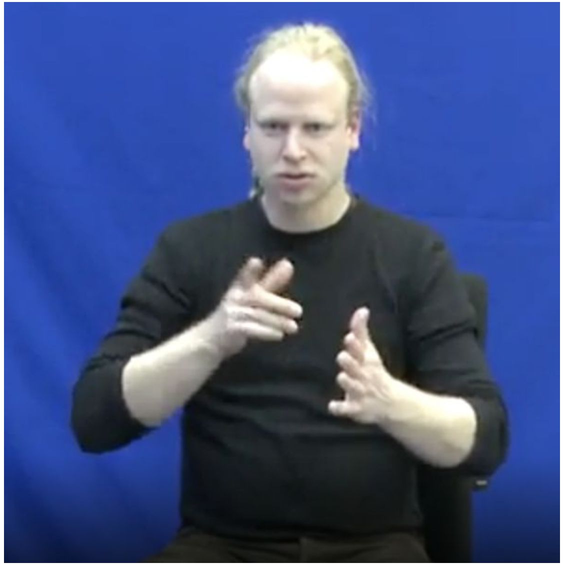


Negative Relations

1. I and I

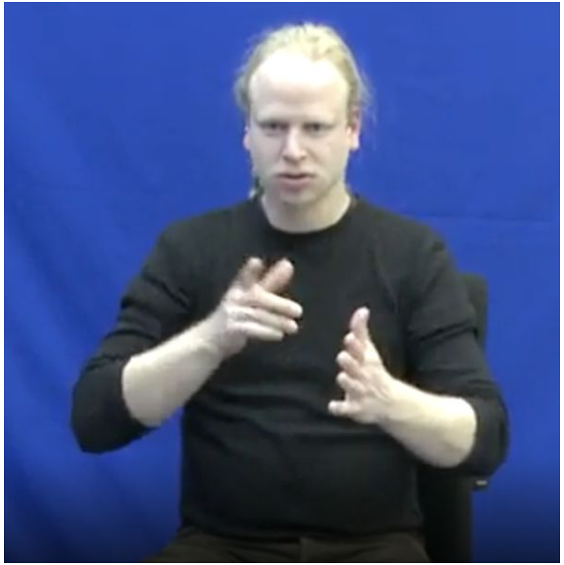
Negative Relations

1. I and I
2. You and You



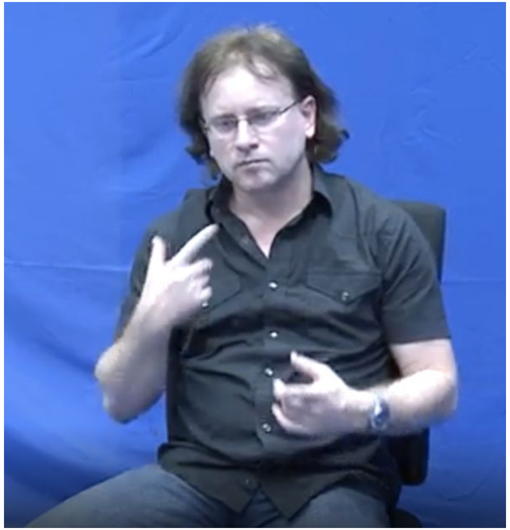
Negative Relations

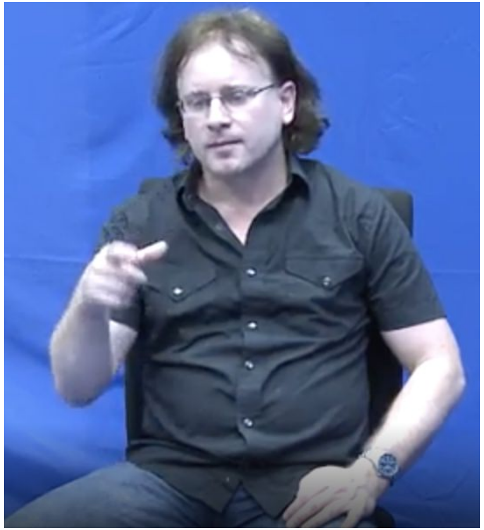
1. I and I
2. You and You
3. I and You



Negative Relations

1. I and I
2. You and You
3. I and You
4. Different Person





Negative Relations

1. I and I
2. You and You
3. I and You
4. Spatially Far Index

Weight Assignment

Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index
5. Noun Phrase
6. Spatially Close Index

Negative Relations

1. I and I
2. You and You
3. I and You
4. Spatially Far Index

Weight Assignment

Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index
5. Noun Phrase
6. Spatially Close Index

Negative Relations

1. I and I
2. You and You
3. I and You
4. Spatially Far Index

$-\infty$

Weight Assignment

Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index
5. Noun Phrase
6. Spatially Close Index

+0.5

Negative Relations

1. I and I
2. You and You
3. I and You
4. Spatially Far Index

$-\infty$

Weight Assignment

Positive Relations

1. I and I
2. You and You
3. I and You
4. Temporally Close Index
5. Noun Phrase
6. Spatially Close Index

+0.5

+(10-t)/20

Negative Relations

1. I and I
2. You and You
3. I and You
4. Spatially Far Index

$-\infty$

Weight Assignment

Positive Relations

1. I and I
2. You and You
3. I and You

+0.5

4. Temporally Close Index
5. Noun Phrase

+(10-t)/20

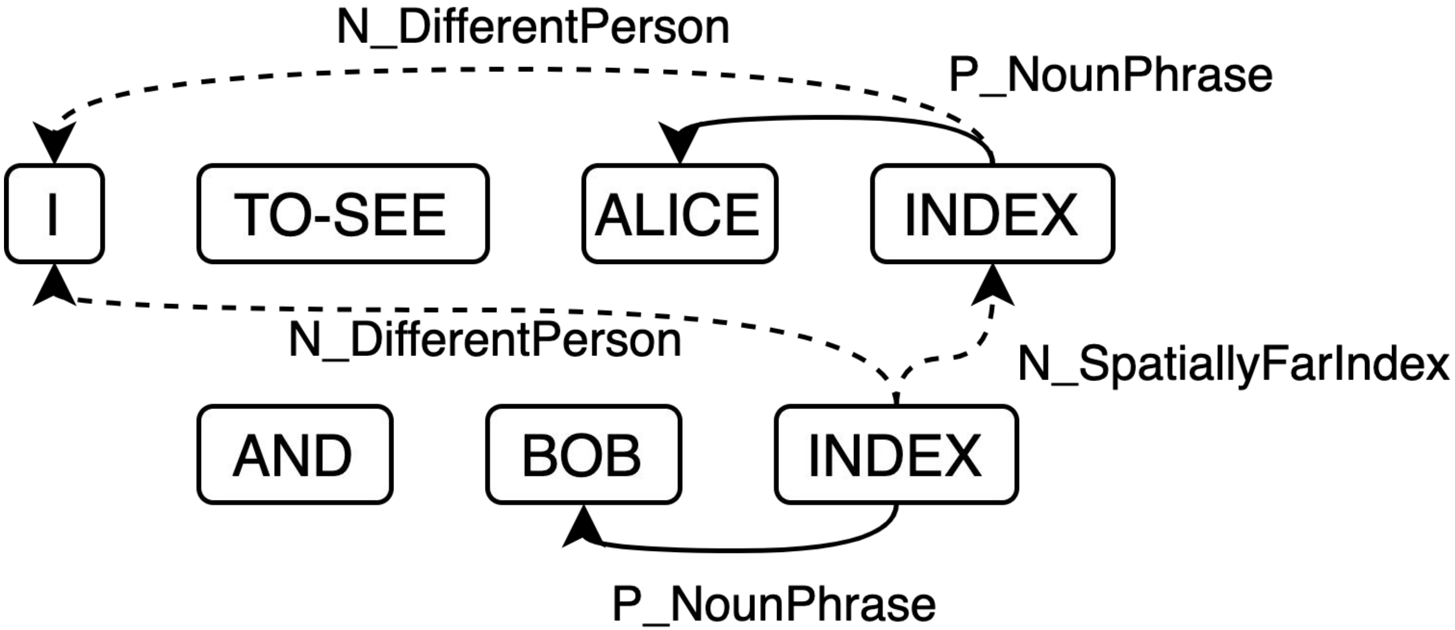
6. Spatially Close Index +(50-s)/50

Negative Relations

1. I and I
2. You and You
3. I and You
4. Spatially Far Index

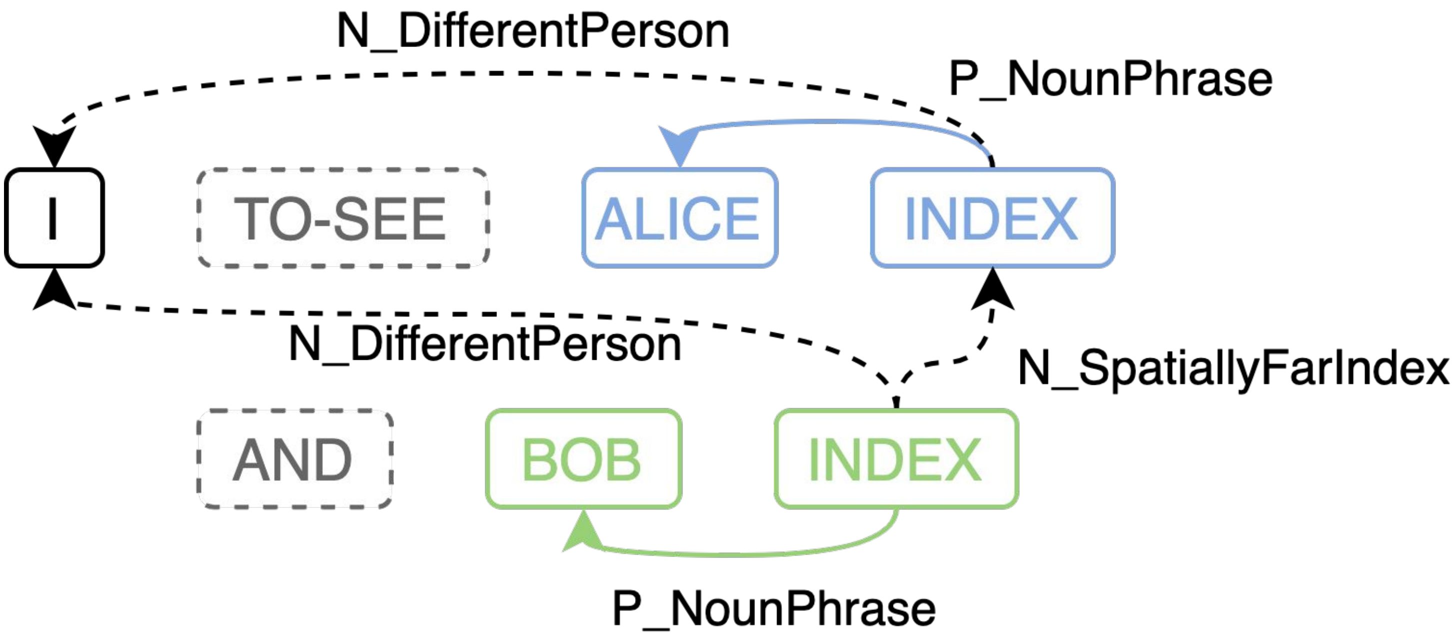
$-\infty$

Clustering



N_DifferentPerson
P_NounPhrase
I
TO-SEE
ALICE
INDEX
N_DifferentPerson
N_SpatiallyFarIndex
AND
BOB
INDEX
P_NounPhrase

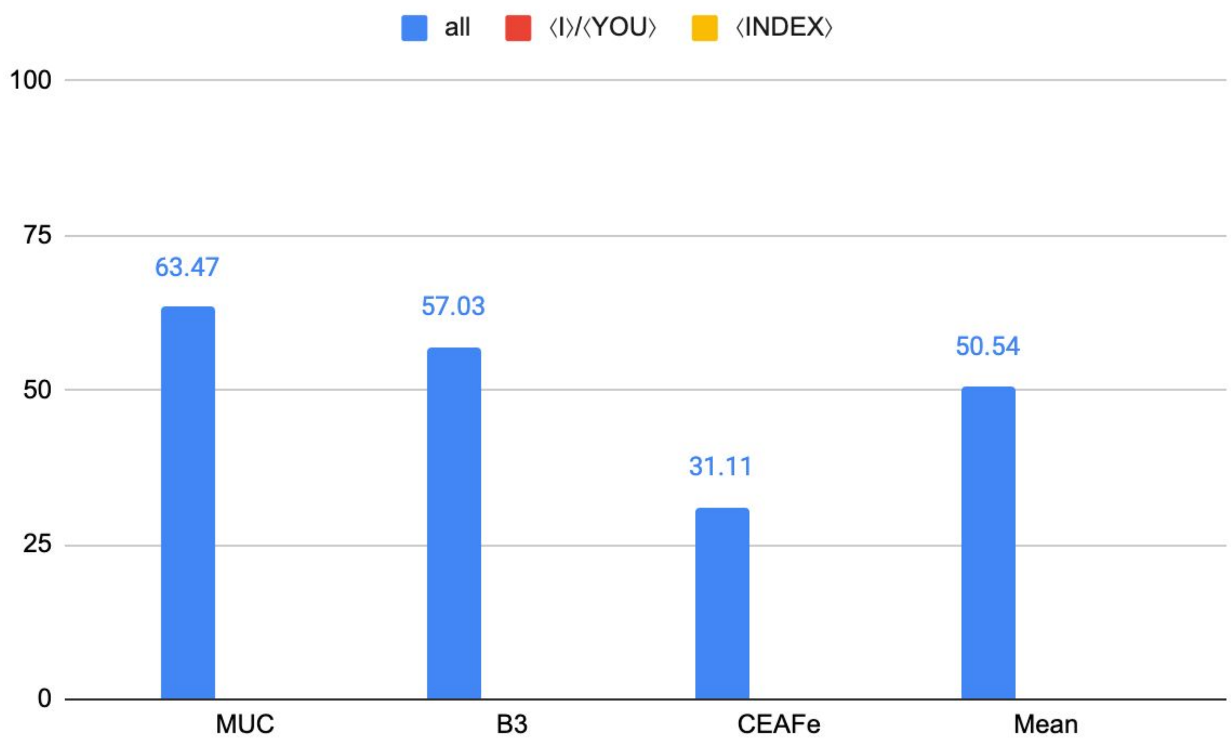
Clustering


N_DifferentPerson
P_NounPhrase
I
TO-SEE
ALICE
INDEX
N_DifferentPerson
N_SpatiallyFarIndex
AND
BOB
INDEX
P_NounPhrase

Outline

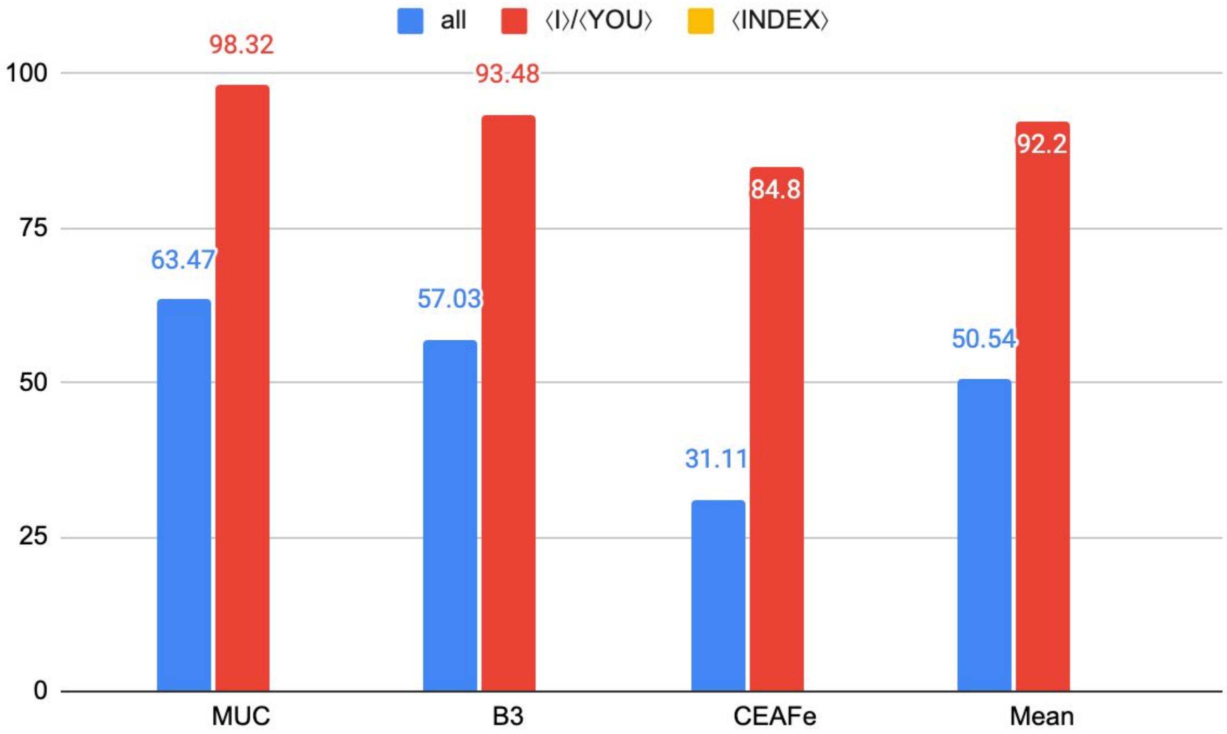
1. Pronominal Pointing Signs
2. Signed Coreference Resolution Task & Data
3. Unsupervised Continuous Multigraph
4. Results & Discussion

Results

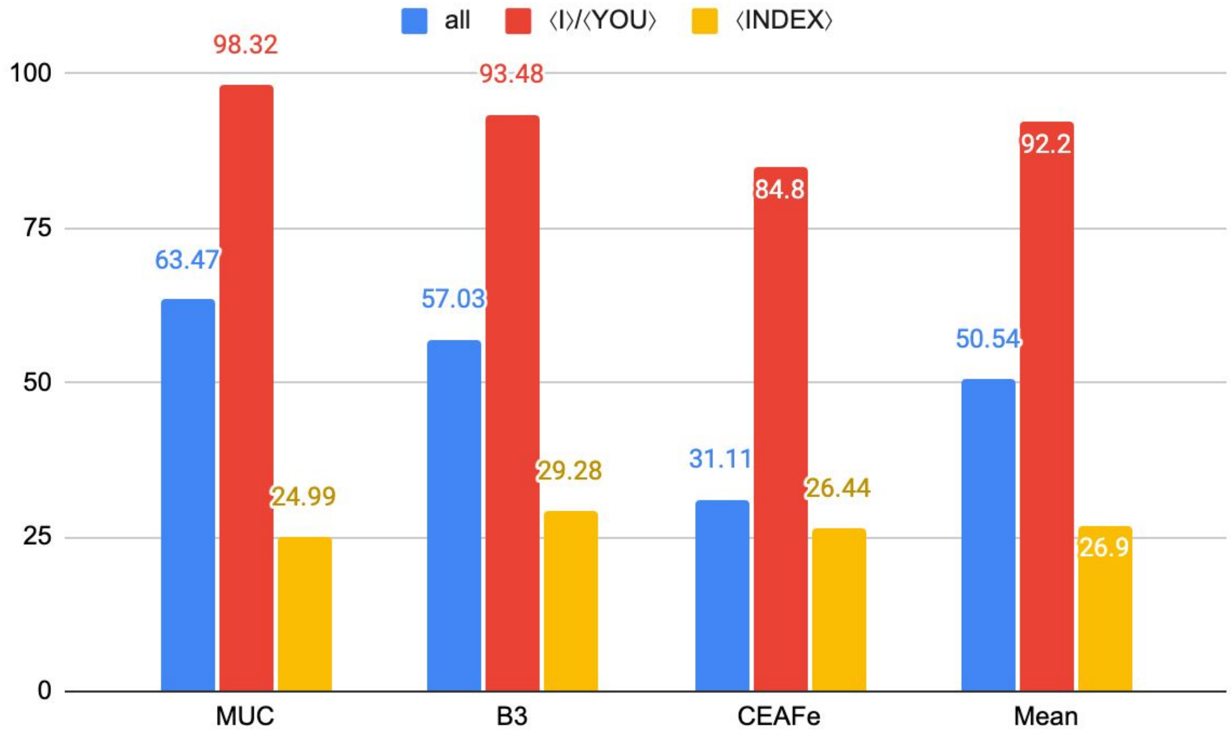


all
⟨I⟩/⟨YOU⟩
⟨INDEX⟩
100
75
50
25
0
63.47
57.03
31.11
50.54
MUC
B3
CEAFe
Mean

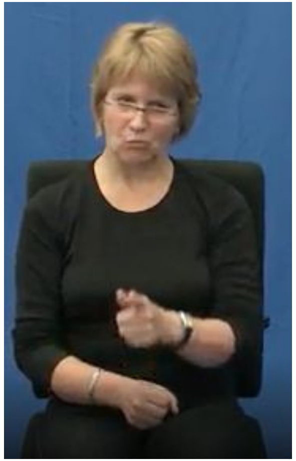
Results


all
⟨I⟩/⟨YOU⟩
⟨INDEX⟩
100
75
50
25
0
98.32
63.47
93.48
57.03
84.8
31.11
92.2
50.54
MUC
B3
CEAFe
Mean

Results

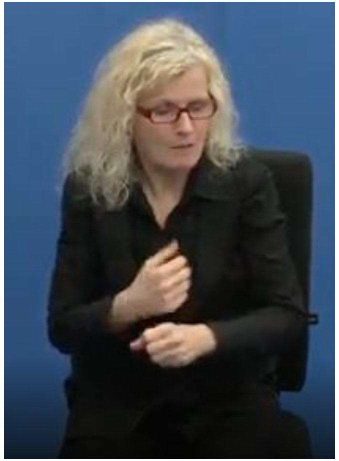

all
⟨I⟩/⟨YOU⟩
⟨INDEX⟩
100
75
50
25
0
63.47
98.32
24.99
57.03
93.48
29.28
31.11
84.8
26.44
50.54
92.2
26.9
MUC
B3
CEAFe
Mean

Examples




TO-SEE **YOU** GOOD **YOU**

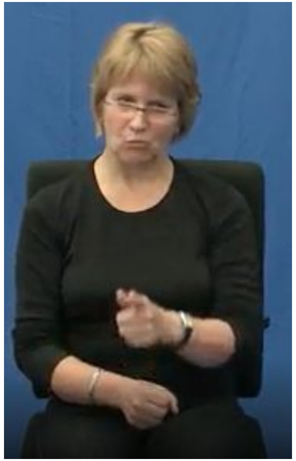
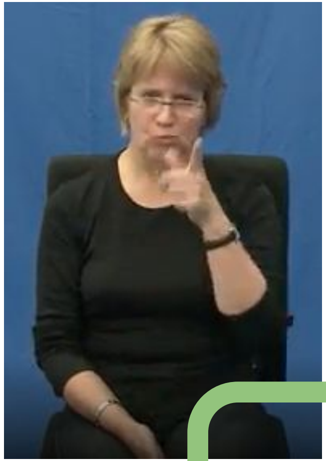
I think you could do a good job there.

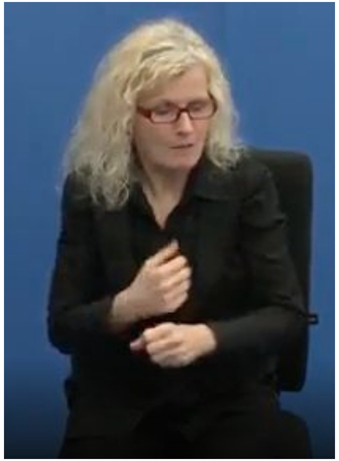



GEST-DECLINE **I** CAN NOT TO-SAY TO-HOLD-ON **I**

I can't keep that promise

Examples



P_IAndYou

TO-SEE **YOU** GOOD **YOU**

I think you could do a good job here.

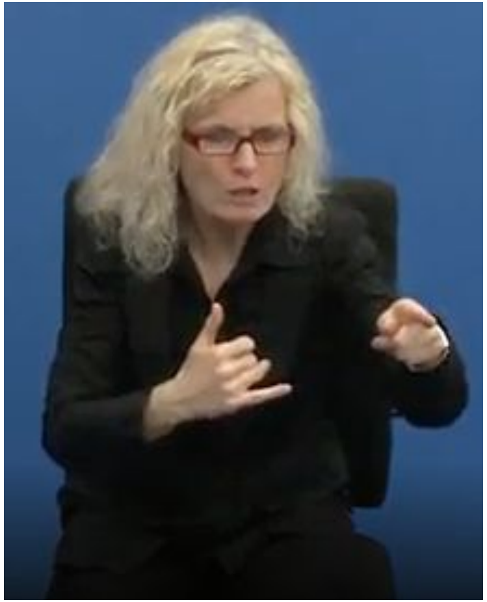
P_YouAndYou

GEST-DECLINE **I** CAN NOT TO-SAY TO-HOLD-ON **I**

I can’t keep that promise

P_IAndI

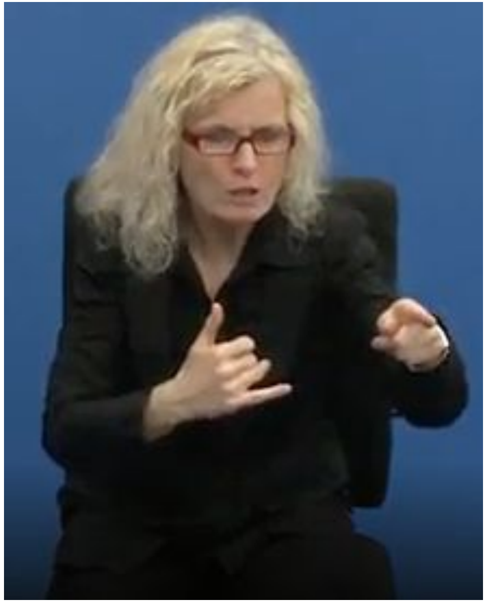
Examples



STUTTGART NUM-1 **NAME INDEX** NUM-1 FREIBURG

Once we were in Stuttgart, once in Ingolstadt and once in Freiburg.

Examples



P_NounPhrase

STUTTGART NUM-1 **NAME INDEX** NUM-1 FREIBURG

Once we were in Stuttgart, once in Ingolstadt and once in Freiburg.

Examples

WITH TRIP **INDEX** SHIP **INDEX**

We went there with an excursion boat.

Examples

P_TemporallyCloseIndex
P_SpatiallyCloseIndex

WITH TRIP **INDEX** SHIP **INDEX**

We went there with an excursion boat.

Examples

I TO-LEARN INDEX **HAMBURG INDEX**

I learned it in Hamburg.

Examples

P_TemporallyCloseIndex
P_SpatiallyCloseIndex

I TO-LEARN **INDEX** **HAMBURG** **INDEX**

I learned it in Hamburg.

Summary

Today's Agenda

➔ Do context-aware machine translation models **pay the right attention**?

➔ When does translation require **context**?

➔ How do we resolve **coreference** in **signed languages**?

Today's Agenda

➔ Do context-aware machine translation models **pay the right attention**?

- ◆ No, but **attention regularization** on **human rationales** can encourage them to do so!

➔ When does translation require **context**?

➔ How do we resolve **coreference** in **signed languages**?

Today's Agenda

- ➔ Do context-aware machine translation models **pay the right attention**?
  - ◆ No, but **attention regularization** on **human rationales** can encourage them to do so!
- ➔ When does translation require **context**?
  - ◆ Ambiguous **pronouns**, lexical **cohesion**, **verb** forms, **formality**, **ellipsis**
- ➔ How do we resolve **coreference** in **signed languages**?

Today's Agenda

- ➔ Do context-aware machine translation models **pay the right attention**?
  - ◆ No, but **attention regularization** on **human rationales** can encourage them to do so!
- ➔ When does translation require **context**?
  - ◆ Ambiguous **pronouns**, lexical **cohesion**, **verb** forms, **formality**, **ellipsis**
- ➔ How do we resolve **coreference** in **signed languages**?
  - ◆ Linguistically-informed **heuristics** and **unsupervised multigraph**