# Why the Time Is Ripe for Discourse in Machine Translation

Rico Sennrich

University of Edinburgh

# A History of Discourse in MT (Abridged)

## Machine Translation will not Work [Kay, 1986]

But, **we still have little idea how to translate** into a closely related language like French or German, English **sentences containing such words as "he", "she", "it"**, "not", "and", and "of". Furthermore, such work as has been done on these problems has been studiously ignored by all those currently involved in developing systems.

# A History of Discourse in MT (Abridged)

## Machine Translation will not Work [Kay, 1986]

But, **we still have little idea how to translate** into a closely related language like French or German, English **sentences containing such words as "he", "she", "it"**, "not", "and", and "of". Furthermore, such work as has been done on these problems has been studiously ignored by all those currently involved in developing systems.

# A History of Discourse in MT (Abridged)

## Anaphora in Rule-based MT

The 1990s have seen an intensification of research efforts in anaphora resolution for MT. This can be seen in the growing number of related projects which have reported promising results (e.g., Wada 1990; Leass & Schwall 1991; Nakaiwa & Ikehara 1992, 1995; Chen 1992; Saggion & Carvalho 1994; Preuss et al. 1994; Nakaiwa et al. 1994, 1995; Mitkov et al. 1995, 1997; Geldbach 1997).

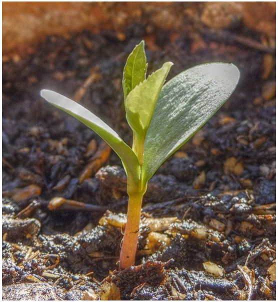
[Mitkov, 1999]

# A History of Discourse in MT (Abridged)

## Anaphora in Rule-based MT

The 1990s have seen an intensification of research efforts in anaphora resolution for MT. This can be seen in the growing number of related projects which have reported promising results (e.g., Wada 1990; Leass & Schwall 1991; Nakaiwa & Ikehara 1992, 1995; Chen 1992; Saggion & Carvalho 1994; Preuss et al. 1994; Nakaiwa et al. 1994, 1995; Mitkov et al. 1995, 1997; Geldbach 1997).

[Mitkov, 1999]



# A History of Discourse in MT (Abridged)

## Discourse in SMT

- anaphora resolution [Le Nagard and Koehn, 2010, Hardmeier and Federico, 2010, Hardmeier et al., 2015]
- lexical consistency [Carpuat, 2009, Tiedemann, 2010, Gong et al., 2011]
- discourse connectives [Meyer et al., 2012]
- topic adaptation [Su et al., 2012, Hasler et al., 2014]

# A History of Discourse in MT (Abridged)

## Discourse in SMT

- anaphora resolution [Le Nagard and Koehn, 2010, Hardmeier and Federico, 2010, Hardmeier et al., 2015]
- lexical consistency [Carpuat, 2009, Tiedemann, 2010, Gong et al., 2011]
- discourse connectives [Meyer et al., 2012]
- topic adaptation [Su et al., 2012, Hasler et al., 2014]

# A History of Discourse in MT (Abridged)

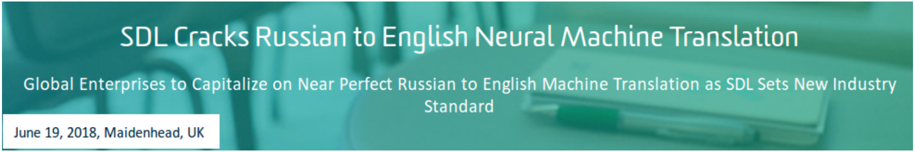
Where Are We Now?

# Where Is Machine Translation Now?

Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn

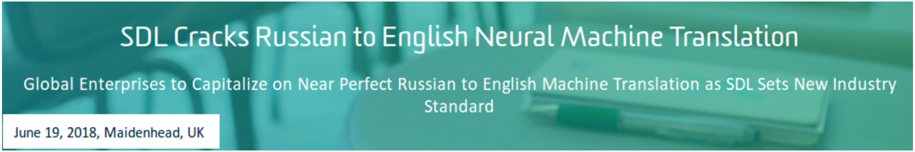


# Where Is Machine Translation Now?

Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn




...extraordinary claims require extraordinary evidence

# Examples from Top WMT18 Systems

coreference

In fairness, Miller did not attack the statue itself.
[...]
But he did attack its meaning [...]

| HUMAN | MT |
|---|---|
| Um fair zu bleiben, Miller griff nicht die Statue selbst an.<br>[...]<br>Aber er griff deren Bedeutung an [...] | Fairerweise hat Miller die Statue nicht selbst angegriffen.<br>[...]<br>Aber er griff seine Bedeutung an [...] |

# Examples from Top WMT18 Systems

coreference

In fairness, Miller did not attack the statue itself.
[...]
But he did attack its meaning [...]

| HUMAN | MT |
|---|---|
| Um fair zu bleiben, Miller griff nicht die Statue selbst an.<br>[...]<br>Aber er griff deren Bedeutung an [...] | Fairerweise hat Miller die Statue nicht selbst angegriffen.<br>[...]<br>Aber er griff seine Bedeutung an [...] |

# Examples from Top WMT18 Systems

lexical coherence

Weidezaunprojekt ist elementar

Das Fischerbacher Weidezaun-Projekt ist ein Erfolgsprojekt und wird im kommenden Jahr fortgesetzt.

| HUMAN | MT |
|---|---|
| Pasture fence project is fundamental | Electric fence project is basic |
| The Fischerbach pasture fence project is a successful project and will be continued next year. | The Fischerbacher Weidezaun-Project is a success and will be continued in the coming year. |

# Examples from Top WMT18 Systems

lexical coherence

Weidezaunprojekt ist elementar

Das Fischerbacher Weidezaun-Projekt ist ein Erfolgsprojekt und wird im kommenden Jahr fortgesetzt.

| HUMAN | MT |
|---|---|
| Pasture fence project is fundamental<br><br>The Fischerbach pasture fence project is a successful project and will be continued next year. | Electric fence project is basic<br><br>The Fischerbacher Weidezaun-Project is a success and will be continued in the coming year. |

# Examples from Top WMT18 Systems

pro-drop

该款机器人使用语音合成、[...]

曾获得国际消费电子产品展（CES）[...]

| HUMAN | MT |
|---|---|
| This robot uses speech synthesis, [...] with conversational [...] features. | Using speech synthesis [...] the robot has the functions of chatting conversation [...] |
| It has won two major CES awards [...] | Has won two awards at the International Consumer Electronics Exhibition (CES) [...] |

# Examples from Top WMT18 Systems

pro-drop

该款机器人使用语音合成、[...]

曾获得国际消费电子产品展（CES）[...]

| HUMAN | MT |
|---|---|
| This robot uses speech synthesis, [...] with conversational [...] features. | Using speech synthesis [...] the robot has the functions of chatting conversation [...] |
| It has won two major CES awards [...] | Has won two awards at the International Consumer Electronics Exhibition (CES) [...] |

# Achieving Human Parity

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn

**laudable...**

- follows best practices with WMT-style evaluation
- data released for scientific scrutiny (outputs, references, rankings)

# Achieving Human Parity

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn

## ...but warrants further scrutiny

- failure to reject null hypothesis is not evidence of parity

# Achieving Human Parity

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn

## ...but warrants further scrutiny

- failure to reject null hypothesis is not evidence of parity
  are we 95% sure that there is a quality difference?
  ...hm...no.
  ~~hey everyone, they're the same! I'm 95% sure!~~

# Achieving Human Parity

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

March 14, 2018 | Allison Linn

## ...but warrants further scrutiny

- failure to reject null hypothesis is not evidence of parity
  are we 95% sure that there is a quality difference?
  ...hm...no.
  ~~hey everyone, they're the same! I'm 95% sure!~~
- alternative hypothesis:
  human raters prefer human translations on a **document-level**
- rationale:
  - context helps raters understand text and spot semantic errors
  - discourse errors are invisible in sentence-level evaluation



can we reproduce Microsoft's finding with different evaluation protocol?

| | original evaluation | our evaluation |
|---|---|---|
| test set | WMT17 | WMT17 (native Chinese part) |
| system | Microsoft COMBO-6 | Microsoft COMBO-6 |
| raters | crowd-workers | **professional translators** |
| experimental unit | sentence | sentence / **document** |
| measurement | direct assessment | **pairwise ranking** |
| raters see reference | no | no |
| raters see source | yes | yes / no |
| ratings | $\geq$ 2,520 per system | $\approx$ 200 per setting |

# Which Text is Better?

Members of the public who find their cars obstructed by unfamiliar vehicles during their daily journeys can use the "Twitter Move Car" feature to address this distress when the driver of the unfamiliar vehicle cannot be reached.

A citizen whose car is obstructed by vehicle and is unable to contact the owner of the obstructing vehicle can use the "WeChat Move the Car" function to address the issue.

# Which Text is Better?

Members of the public who find their cars obstructed by unfamiliar vehicles during their daily journeys can use the "Twitter Move Car" feature to address this distress when the driver of the unfamiliar vehicle cannot be reached.

On August 11, Xi'an traffic police WeChat service number "Xi'an traffic police" launched "WeChat mobile" service.

With the launch of the service, members of the public can tackle such problems in their daily lives by using the "WeChat Move" feature when an unfamiliar vehicle obstructs the movement of their vehicle while the driver is not at the scene. [...]

A citizen whose car is obstructed by vehicle and is unable to contact the owner of the obstructing vehicle can use the "WeChat Move the Car" function to address the issue.

The Xi'an Traffic Police WeChat official account "Xi'an Jiaojing" released the "WeChat Move the Car" service since August 11.

Once the service was released, a fellow citizen whose car was obstructed by another vehicle and where the driver of the vehicle was not present, the citizen could use the "WeChat Move the Car" function to address the issue. [...]

# Which Text is Better?

市民在日常出行中,发现爱车被陌生车辆阻碍了,在联系不上陌生车辆司机的情况下,可以使用“微信挪车”功能解决这一困扰。

8月11日起,西安交警微信服务号“西安交警”推出“微信挪车”服务。

这项服务推出后,日常生活中,市民如遇陌生车辆在驾驶人不在现场的情况下阻碍自己车辆行驶时,就可通过使用“微信挪车”功能解决此类问题。[...]

Members of the public who find their cars obstructed by unfamiliar vehicles during their daily journeys can use the "Twitter Move Car" feature to address this distress when the driver of the unfamiliar vehicle cannot be reached.

On August 11, Xi'an traffic police WeChat service number "Xi'an traffic police" launched "WeChat mobile" service.

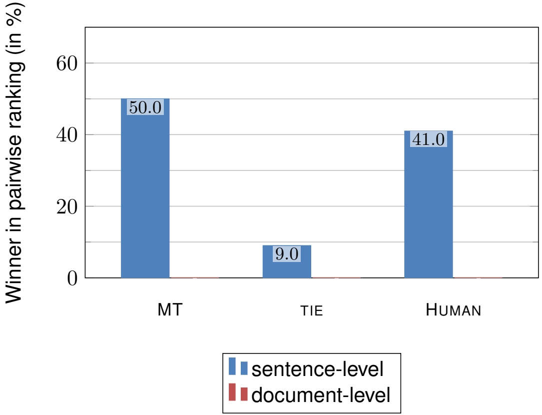
With the launch of the service, members of the public can tackle such problems in their daily lives by using the "WeChat Move" feature when an unfamiliar vehicle obstructs the movement of their vehicle while the driver is not at the scene. [...]

A citizen whose car is obstructed by vehicle and is unable to contact the owner of the obstructing vehicle can use the "WeChat Move the Car" function to address the issue.

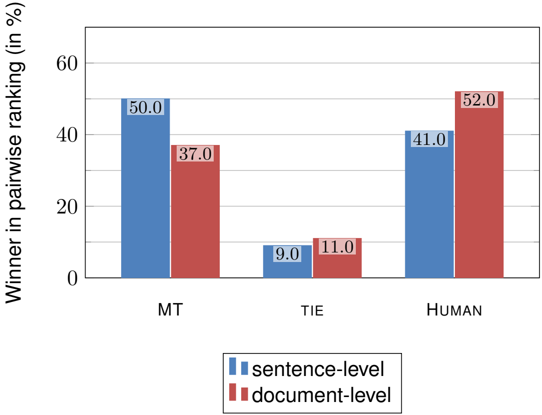
The Xi'an Traffic Police WeChat official account "Xi'an Jiaojing" released the "WeChat Move the Car" service since August 11.

Once the service was released, a fellow citizen whose car was obstructed by another vehicle and where the driver of the vehicle was not present, the citizen could use the "WeChat Move the Car" function to address the issue. [...]

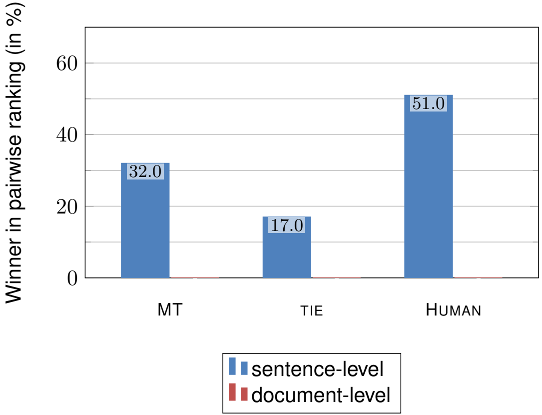
# Evaluation Results: Bilingual Assessment



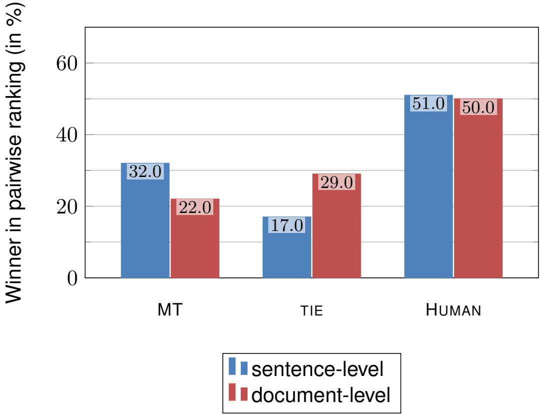
# Evaluation Results: Bilingual Assessment



# Evaluation Results: Monolingual Assessment



# Evaluation Results: Monolingual Assessment



# A Case for Document-level Evaluation

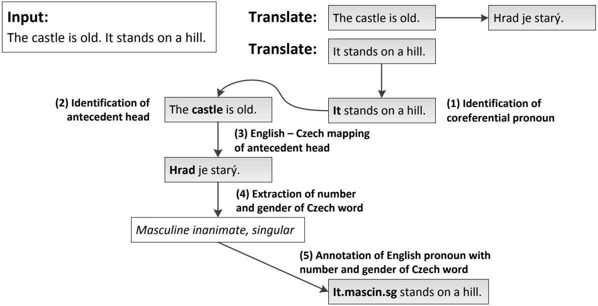
- document-level ratings show significant preference for HUMAN
- preference for HUMAN is even stronger in monolingual evaluation

## Conclusions

- discourse-level cohesion and coherence is important, but invisible in sentence-level evaluation
- distinguishing MT from human translations becomes harder with increasing quality
  $\rightarrow$ is it time to move to document-level evaluation in shared tasks?

# New Chances: Context-Aware NMT

SMT era:

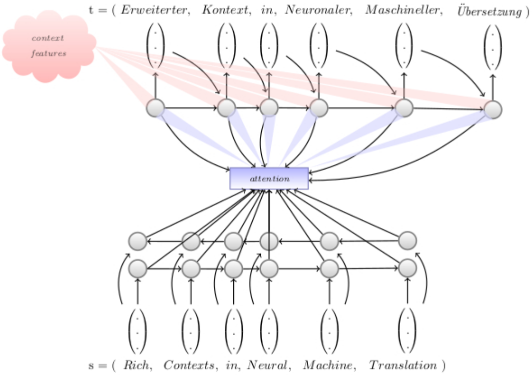



specialized features

[Hardmeier, 2012, Guillou, 2012, Meyer et al., 2012]

# New Chances: Context-Aware NMT

NMT era:




contextual sentences as additional input

[Jean et al., 2017, Wang et al., 2017, Tiedemann and Scherrer, 2017, Bawden et al., 2018, Voita et al., 2018, Maruf and Haffari, 2018]

# Some Research Questions

- How do we measure progress?
- Which context matters?
- What neural architectures work well?

# Evaluating Discourse Phenomena

[Bawden et al., NAACL 2018]




- How do we measure progress?
  $\rightarrow$ hand-crafted test set of 200 context-dependent translations
- Which context matters?
  $\rightarrow$ (focus on translations that depend on previous target sentence)
- What neural architectures work well?
  $\rightarrow$ exploration of multi-encoder and concatenating architectures

setup: train on subset of OpenSubtitles2016 English-French

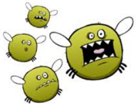
# A Contrastive Test Set: Coreference

**Source:**

| | |
|---|---|
| context: | Oh, I hate **flies**. Look, there's another one! |
| sentence: | Don't worry, I'll kill **it** for you. |

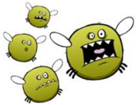
**Target:**

| | |
|---|---|
| context: | Ô je déteste les mouches.<br>Regarde, il y en a une autre ! |
| correct: | T'inquiète, je **la** tuerai pour toi. |
| incorrect: | T'inquiète, je **le** tuerai pour toi. |



# A Contrastive Test Set: Coreference

Can the model rank the correct sentence above the incorrect one?



**Source:**
context: Oh, I hate **flies**. Look, there's another one!
sentence: Don’t worry, I'll kill **it** for you.

**Target:**
context: Ô je déteste les mouches.
Regarde, il y en a une autre !
correct: T'inquiète, je **la** tuerai pour toi.
incorrect: T'inquiète, je **le** tuerai pour toi.

# A Contrastive Test Set: Coreference

Can the model rank the correct sentence above the incorrect one?

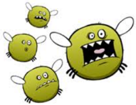
Previous linguistic context necessary to disambiguate

**Source:**

| | |
|---|---|
| context: | Oh, I hate **flies**. Look, there's another one! |
| sentence: | Don’t worry, I'll kill **it** for you. |

**Target:**

| | |
|---|---|
| context: | Ô je déteste les mouches. Regarde, il y en a une autre ! |
| correct: | T'inquiète, je **la** tuerai pour toi. |
| incorrect: | T'inquiète, je **le** tuerai pour toi. |



# A Contrastive Test Set: Coreference

Can the model rank the correct sentence above the incorrect one?

Previous linguistic context necessary to disambiguate

**Source:**

context: Oh, I hate **flies**. Look, there's another one!

sentence: Don’t worry, I'll kill **it** for you.

**Target:**

context: Ô je déteste les mouches.
Regarde, il y en a une autre !

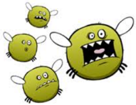
correct: T'inquiète, je **la** tuerai pour toi.

incorrect: T'inquiète, je **le** tuerai pour toi.

context: Ô je déteste les **moucherons**.
Regarde, il y en a un autre !

correct: T'inquiète, je **le** tuerai pour toi.

incorrect: T'inquiète, je **la** tuerai pour toi.



**Balanced examples:** Non-contextual baseline scores 50%

# A Contrastive Test Set: Coherence and Cohesion

| | |
|---|---|
| **Source:** | |
| context: | So what do you say to £50? |
| current sent.: | It's a little **steeper** than I was expecting. |
| **Target:** | |
| context: | Qu'est-ce que vous en pensez de 50£ ? |
| correct: | C'est un peu plus **cher** que ce que je pensais. |
| incorrect: | C'est un peu plus **raide** que ce que je pensais. |

| | |
|---|---|
| **Source:** | |
| context: | How are your feet holding up? |
| current sent.: | It's a little **steeper** than I was expecting. |
| **Target:** | |
| context: | Comment vont tes pieds ? |
| correct: | C'est un peu plus **raide** que ce que je pensais. |
| incorrect: | C'est un peu plus **cher** que ce que je pensais. |

# A Contrastive Test Set: Coherence and Cohesion

**Source:**

| | |
|---|---|
| context: | What's **crazy** about me? |
| current sent.: | Is this **crazy**? |

**Target:**

| | |
|---|---|
| context: | Qu'est-ce qu'il y a de **dingue** chez moi ? |
| correct: | Est-ce que ça c'est **dingue** ? |
| incorrect: | Est-ce que ça c'est fou ? |

---

**Source:**

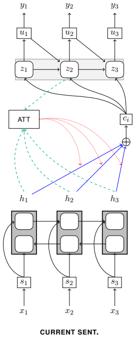
| | |
|---|---|
| context: | What's **crazy** about me? |
| current sent.: | Is this **crazy**? |

**Target:**

| | |
|---|---|
| context: | Qu'est-ce qu'il y a de **fou** chez moi ? |
| correct: | Est-ce que ça c'est **fou** ? |
| incorrect: | Est-ce que ça c'est dingue ? |

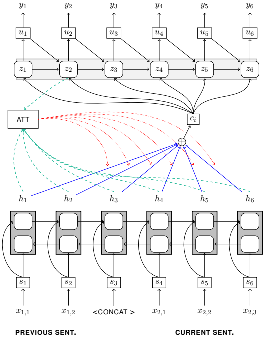
# Architectures

## Baseline



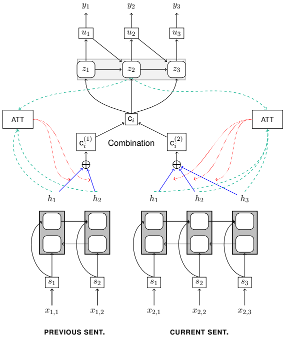

[Bahdanau et al., 2015]

## 2TO2 - concatenated input




[Tiedemann and Scherrer, 2017]

## Multiple encoders



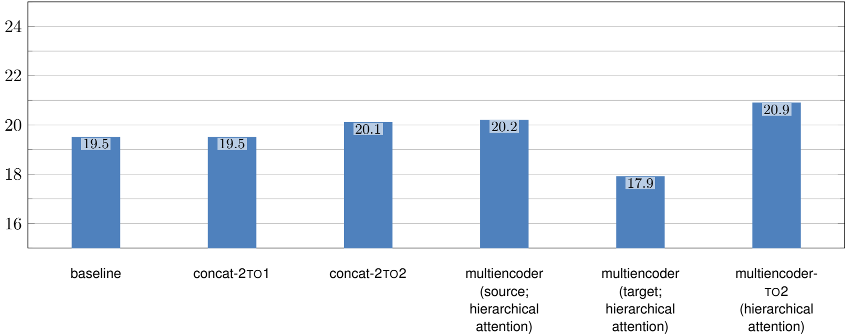

[Wang et al., 2017]

# Architectures

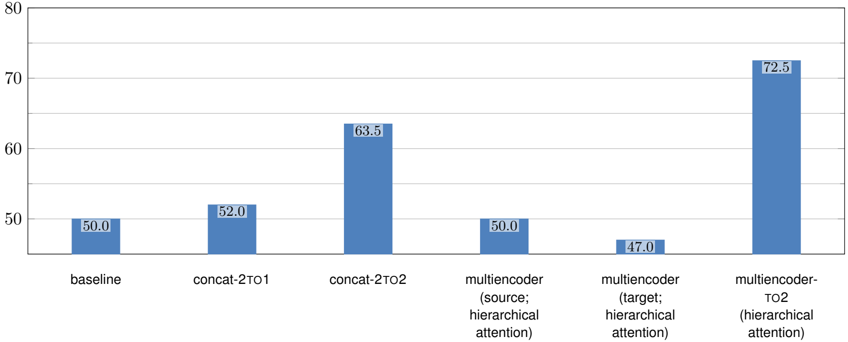
architecture exploration:

- condition on previous source, target, or both?
- use multiple encoders or just concatenate sentences?
- how to combine multiple context vectors in multi-encoder setups?
    - concatenate
    - gating mechanism
    - hierarchical attention

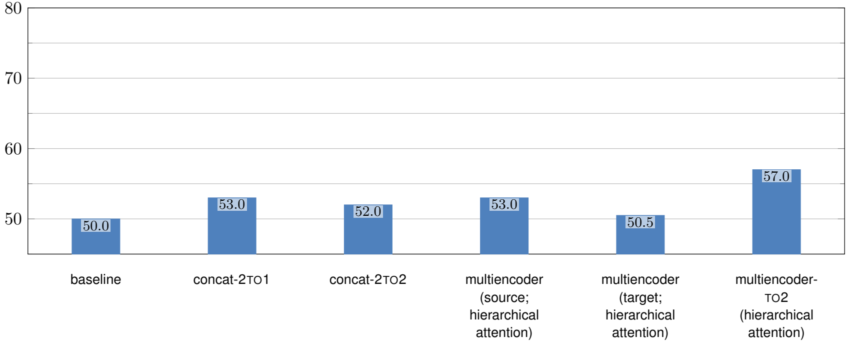
# Results: BLEU



# Results: Contrastive Test Set: Coreference



# Results: Contrastive Test Set: Coherence/Cohesion



# Evaluating Discourse Phenomena: Conclusions

- simple context-aware NMT systems learn discourse phenomena
- architectures matter
- learning coreference easier than lexical coherence and
  word sense disambiguation (?)
  (for hand-crafted, difficult cases)

- future work: more comprehensive test set of discourse phenomena

# Analyzing Use of Context

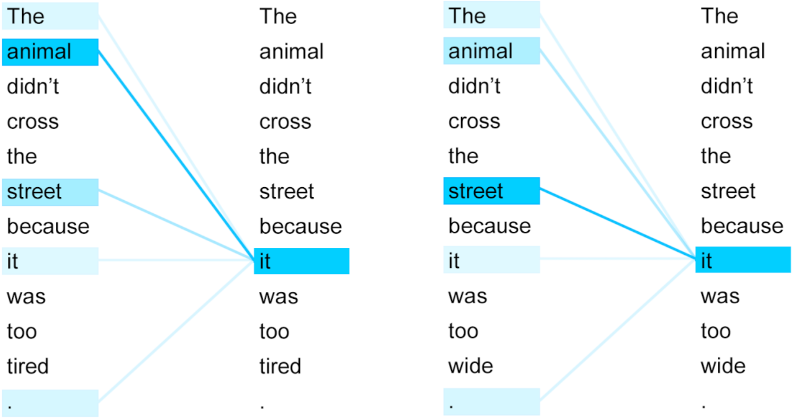
[Tiedemann and Scherrer, 2017]


### set-up

- simple architecture: concatenation of previous sentence
- analysis of attention patterns
- recurrent connections make analysis difficult


Figure 6: Attention patterns with referential pronouns in extended context.

# Analyzing Use of Context: Transformer




[Uszkoreit, 2017]

# Analyzing Use of Context

[Voita, Serdyukov, Sennrich, Titov, ACL 2018]




**set-up**

- Transformer architecture with clear interface to context
- analysis of attention patterns

# Context-Aware Transformer

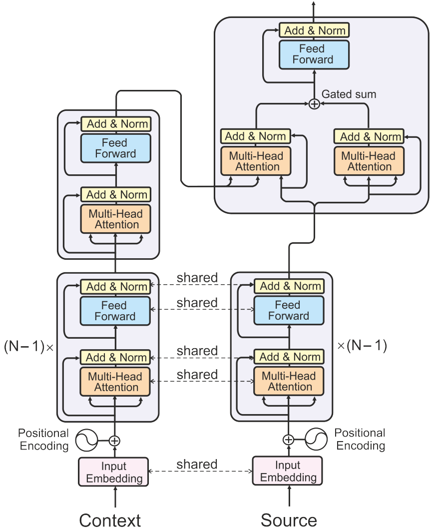
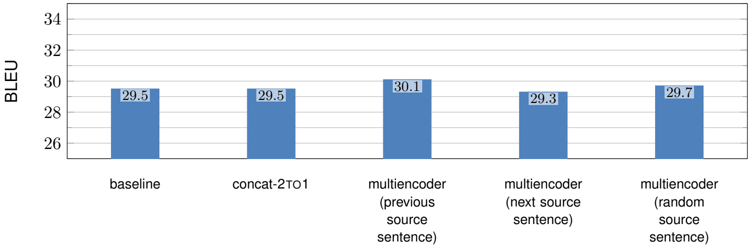


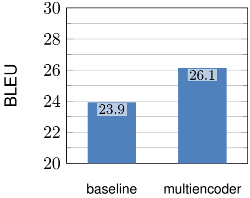

Figure 1: Encoder of the discourse-aware model

# Context-Aware Transformer: Evaluation

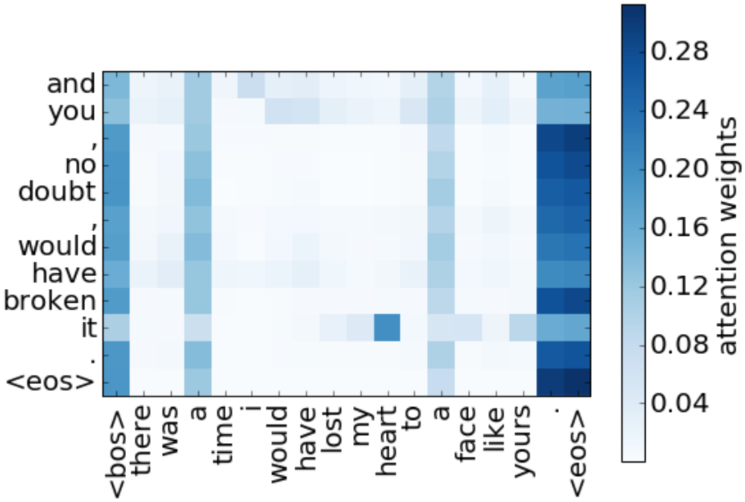
- OpenSubtitles2018 English→Russian
- scores on random test set:




larger improvements on focused test set ('it' with nominal antecedent):



# Context-Aware Transformer Learns Anaphora Resolution



# Context-Aware Transformer Learns Anaphora Resolution

| | agreement (in %) |
|---|---|
| coreNLP | 77 |
| attention | 72 |
| last noun | 54 |

Agreement with human assessment for coreference resolution of anaphoric *it*.
Examples with $\geq 1$ noun in context sentence.

# Discourse in MT

# How Do We Measure Progress?

- BLEU still works (somewhat)
- targeted evaluation:
  - contrastive pairs
  - pronoun translation task
- document-level human evaluation

potential for new discourse-level measures

# Which Context Matters?

most work so far focuses on previous sentence, but:

- relevant information can be further in past
- relevant information can be in future context

| | |
|---|---|
| source | I went there with **my friend**.<br>**She** was amazed to see that it had multiple floors. |
| reference | Sono andato la' con **la mia amica**.<br>E' rimasta meraviglia nel vedere che aveva piu' piani |
| baseline | Arrivai li con **il mio amico**.<br>Rimaneva meravigliato di vedere che aveva una cosa piu incredibile. |
| contextual (prev+next) | Sono andato con **la mia amica**.<br>Fu sorpresa nel vedere che aveva piu piani. |

[Agrawal et al., 2018]

# What Neural Architectures Work Well?

- baseline architectures ok(-ish) with concatenated context
- simple multi-encoder architectures effective

big challenge: efficiently/effectively scaling up to large contexts

# Conclusions

- we cannot "solve" machine translation on sentence-level
- let's put effort into:
  - document-level training data
  - document-level evaluation
- simple context-aware architectures learn discourse phenomena...
  ...but still work to be done on better architectures for large contexts

# Conclusions

- we cannot "solve" machine translation on sentence-level
- let's put effort into:
  - document-level training data
  - document-level evaluation
- simple context-aware architectures learn discourse phenomena...
  ...but still work to be done on better architectures for large contexts

**Thank you for your attention**

## Resources

- contrastive test sets for discourse in MT evaluation:
  `https://github.com/rbawden/discourse-mt-test-sets`

# Bibliography I


Agrawal, R., Turchi, M., and Negri, M. (2018).
Contextual Handling in Neural Machine Translation: Look Behind, Ahead and on Both Sides.
In 21th Annual Conference of the European Association for Machine Translation.

Bahdanau, D., Cho, K., and Bengio, Y. (2015).
Neural Machine Translation by Jointly Learning to Align and Translate.
In Proceedings of the International Conference on Learning Representations (ICLR).

Bawden, R., Sennrich, R., Birch, A., and Haddow, B. (2018).
Evaluating Discourse Phenomena in Neural Machine Translation.
In NAACL 2018, New Orleans, USA.

Carpuat, M. (2009).
One Translation Per Discourse.
In Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions, pages 19–27, Boulder, Colorado. Association for Computational Linguistics.

Gong, Z., Zhang, M., and Zhou, G. (2011).
Cache-based Document-level Statistical Machine Translation.
In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pages 909–919, Edinburgh, Scotland, UK. Association for Computational Linguistics.

Guillou, L. (2012).
Improving Pronoun Translation for Statistical Machine Translation.
In
Proceedings of the Student Research Workshop at the 13th Conference of the European Chapter of the Association for Computation
pages 1–10, Avignon, France. Association for Computational Linguistics.

# Bibliography II


Hardmeier, C. (2012).
Discourse in statistical machine translation: A survey and a case study.
Discours, 11.

Hardmeier, C. and Federico, M. (2010).
Modelling Pronominal Anaphora in Statistical Machine Translation.
In Federico, M., Lane, I., Paul, M., and Yvon, F., editors,
Proceedings of the seventh International Workshop on Spoken Language Translation (IWSLT), pages 283–289.

Hardmeier, C., Nakov, P., Stymne, S., Tiedemann, J., Versley, Y., and Cettolo, M. (2015).
Pronoun-Focused MT and Cross-Lingual Pronoun Prediction: Findings of the 2015 DiscoMT Shared Task on Pronoun Translation.
In Proceedings of the Second Workshop on Discourse in Machine Translation, pages 1–16, Lisbon, Portugal. Association for Computational Linguistics.

Hasler, E., Blunsom, P., Koehn, P., and Haddow, B. (2014).
Dynamic Topic Adaptation for Phrase-based MT.
In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pages 328–337, Gothenburg, Sweden.

Jean, S., Lauly, S., Firat, O., and Cho, K. (2017).
Neural Machine Translation for Cross-Lingual Pronoun Prediction.
In Proceedings of the 3rd Workshop on Discourse in Machine Translation, DISCOMT'17, pages 54–57, Copenhagen, Denmark.

Kay, M. (1986).
Machine Translation will not Work.
In Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics, pages 268–268, New York, New York, USA. Association for Computational Linguistics.

# Bibliography III


Le Nagard, R. and Koehn, P. (2010).
Aiding Pronoun Translation with Co-Reference Resolution.
In Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR, pages 252–261, Uppsala, Sweden. Association for Computational Linguistics.

Maruf, S. and Haffari, G. (2018).
Document Context Neural Machine Translation with Memory Networks.
In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1275–1284. Association for Computational Linguistics.

Meyer, T., Popescu-Belis, A., Hajlaoui, N., and Gesmundo, A. (2012).
Machine Translation of Labeled Discourse Connectives.
In Proceedings of the Tenth Conference of the Association for Machine Translation in the Americas (AMTA).

Mitkov, R. (1999).
Introduction: Special Issue on Anaphora Resolution in Machine Translation and Multilingual NLP.
Machine Translation, 14(3/4):159–161.

Su, J., Wu, H., Wang, H., Chen, Y., Shi, X., Dong, H., and Liu, Q. (2012).
Translation Model Adaptation for Statistical Machine Translation with Monolingual Topic Information.
In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 459–468, Jeju Island, Korea. Association for Computational Linguistics.

Tiedemann, J. (2010).
Context Adaptation in Statistical Machine Translation Using Models with Exponentially Decaying Cache.
In Proc. ACL Workshop on Domain Adaptation for Natural Language Processing, pages 8–15, Uppsala, Sweden. Association for Computational Linguistics.

# Bibliography IV


Tiedemann, J. and Scherrer, Y. (2017).
Neural Machine Translation with Extended Context.
In Proceedings of the Third Workshop on Discourse in Machine Translation, pages 82–92, Copenhagen, Denmark. Association for Computational Linguistics.

Uszkoreit, J. (2017).
Transformer: A Novel Neural Network Architecture for Language Understanding.
https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html.

Voita, E., Serdyukov, P., Sennrich, R., and Titov, I. (2018).
Context-Aware Neural Machine Translation Learns Anaphora Resolution.
In ACL 2018, Melbourne, Australia.

Wang, L., Tu, Z., Way, A., and Qun Liu (2017).
Exploiting Cross-Sentence Context for Neural Machine Translation.
In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP'17, pages 2816–2821, Denmark, Copenhagen.