Document Simplification

Controlling Simplification Operations and
Modeling Document Context

Claire Gardent

Joint work with Liam Cripwell

CNRS / LORIA, Nancy, France

What is Document Simplification?

Complex Input Document

Owls are birds from the order of Strigiformes, comprising over 200 species of mostly) solitary and nocturnal birds of prey typified by an upright stance, binocular vision, binaural hearing, and sharp talons. Owls hunt mostly small mammals, insects, and other birds, although a few species specialize in hunting fish.

Simplified Output Document

Owls are birds. There are over 200 species and are all animals of prey. Most of them are solitary and nocturnal. Owls' prey may be birds, large insects (such as crickets), small reptiles (such as lizards) or small mammals (such as mice, rats, and rabbits).

Avg nb of sentences in Input Document: 39

Why Simplify ?

To aid reader comprehension (Mason, 1978; Williams et al., 2003; Kajiwara et al., 2013)

- Adult vs children
- Native vs non Native
- Reading disability
- Expert vs non-Expert

Reading Ability

Simplification Operations

Owls are birds from the order of Strigiformes, comprising over 200 species of mostly solitary and nocturnal birds of prey typified by an upright stance, binocular vision, binaural hearing, and sharp talons. Owls hunt mostly small mammals, insects, and other birds, although a few species specialize in hunting fish.

Owls are birds. There are over 200 species ...

Splitting

Document Simplification

Owls are birds from the order of Strigiformes, comprising over 200 species of ***mostly solitary and nocturnal birds of prey*** typified by an upright stance, binocular vision, binaural hearing, and sharp talons. Owls hunt mostly small mammals, insects, and other birds, although a few species specialize in hunting fish.

Owls are birds. There are over 200 species ***and are all animals of prey. Most of them are solitary and nocturnal*** . Owls' prey may be birds, large insects (such as crickets), small reptiles (such as lizards) or small mammals (such as mice, rats, and rabbits).

Rephrasing

Document Simplification

Owls are birds from the order of Strigiformes, comprising over 200 species of mostly solitary and nocturnal birds of prey ***typified by an upright stance, binocular vision, binaural hearing, and sharp talons*** . Owls hunt mostly small mammals, insects, and other birds, although a few species specialize in hunting fish.

Owls are birds. There are over 200 species and are all animals of prey. Most of them are solitary and nocturnal. Owls’ prey may be birds, large insects (such as crickets), small reptiles (such as lizards) or small mammals (such as mice, rats, and rabbits).

Deleting

Previous work

Sentence-level simplification iteratively applied over a document (Woodsend and Lapata, 2011a; Alva-Manchego et al., 2019b)

Low discourse coherence

Previous work

Sentence-level simplification iteratively applied over a document
(Woodsend and Lapata, 2011a; Alva-Manchego et al., 2019b)

Low discourse coherence

A sentence-level model that uses context information to influence document simplification
(Sun et al. 2020)

Underperform the iterative sentence simplification baseline

Our Model: two Key Components

Planning

PLAN - A sequence of simplification operations for the input document

Our Model: two Key Components

Planning

PLAN - A sequence of simplification operations for the input document

Modeling Context

Simplification operations are predicted based on local and global context

LOCAL - The words making up a sentence

GLOBAL - The text surrounding a sentence

Outline

Planning

$c1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$

Outline

Planning

$c1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$

Simplifying

Plan-guided Document Simplification

Document context is used to predict simplification operations

Outline

Planning

$c_1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$

Simplifying

Plan-guided Document Simplification

Document context is used to predict simplification operations

Context-aware and Plan-guided Document Simplification

Document context is also used to guide simplification

Planning Simplification Operations

$$c1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$$

Cripwell et al. Findings of NAACL 2022

Planning Simplifications

$$c_1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$$

with $\hat{o}_i \in \{$***copy, rephrase, split, delete***$\}$

Given some input document $C = c_1, \ldots, c_n$ the task of the planner is to predict a ***simplification plan*** i.e., a sequence of n simplification operations

$$PLAN = \hat{o}, \ldots, \hat{o}_n$$

Challenges

Simplification Operations have different requirements

Splitting

- mainly depends on the ***input sentence's internal structure***

  The man ***who*** *sleeps snores* → *The man sleeps. He snores.*

  John went shopping ***after*** *he left work* → *John left work. Afterwards he went shopping.*

Challenges

Simplification Operations have different requirements

Splitting

- mainly depends on the ***input sentence's internal structure***

  *The man **who** sleeps snores* → *The man sleeps. He snores.*

  *John went shopping **after** he left work* → *John left work. Afterwards he went shopping.*

Deletion, copy and rephrase

- are mostly ***context dependent*** .

  A sentence can only be omitted if it is either **redundant** with, or of **minor semantic import** relative to, other sentences in the document

Challenges

Simplification Operations have different requirements

Splitting

- mainly depends on the ***input sentence's internal structure***

  ⇒ we model complex sentences at the token level

  LOCAL context

Deletion, copy and rephrase

- are mostly ***context dependent*** .

  ⇒ we take into account the document context of the complex sentences

  GLOBAL context

Planning Model

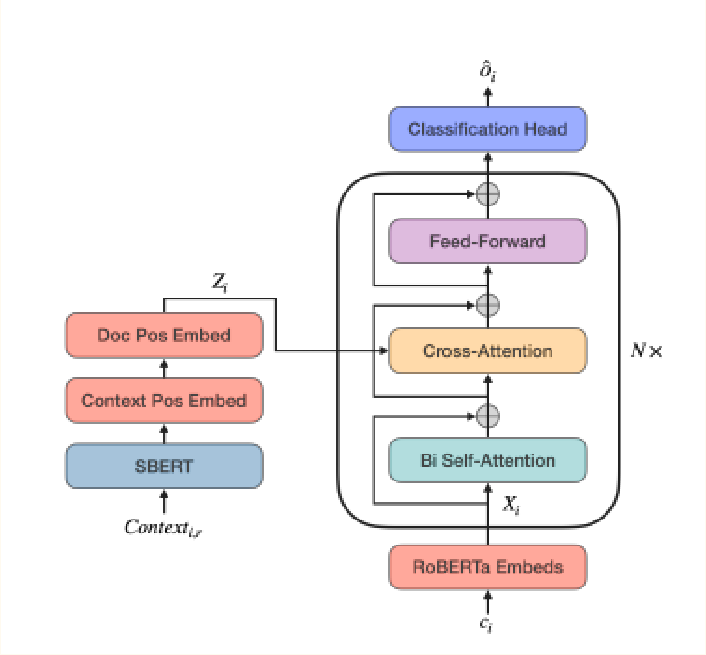
RoBERTa classifier with cross-attention over the global context

Local Context

- ***Token level*** encoder of the sentence to be simplified c_i

Global Context

- fixed window of Sentence level embedding (SBERT) for ***surrounding sentences***



Planning Model

RoBERTa classifier with cross-attention over the global context

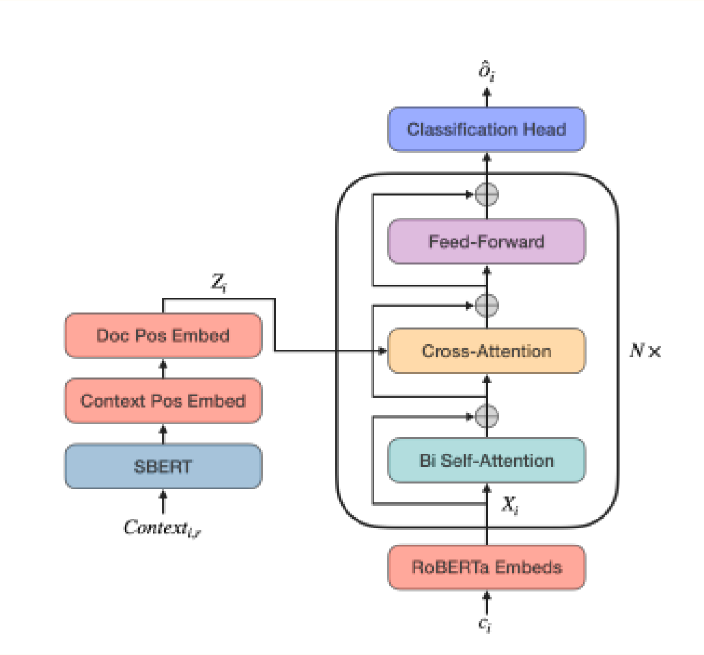
- *layers initialised with weights from a context-independent classifier*

Local Context

- ***Token level*** encoder of the sentence to be simplified c_i

Global Context

- fixed window of Sentence level embedding (SBERT) for ***surrounding sentences***



Planning Model

RoBERTa classifier with cross-attention over the global context

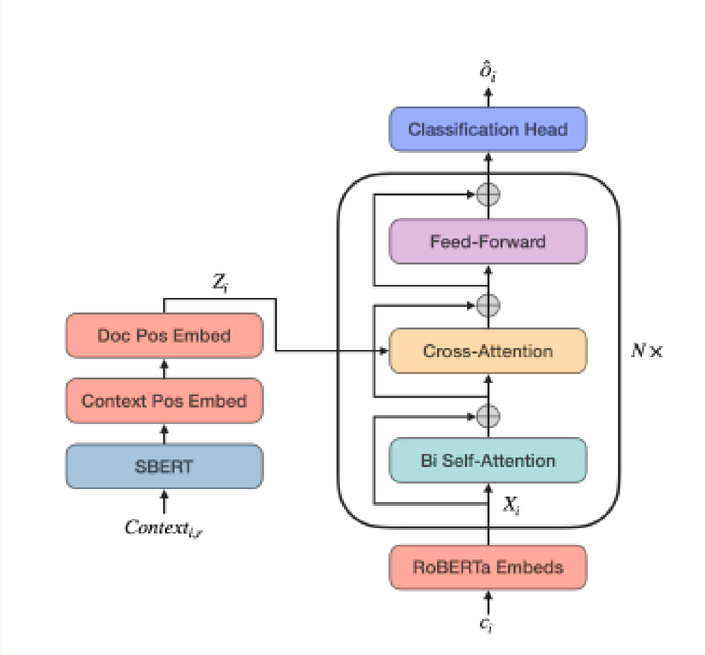
- ***layers initialised with weights from a context-independent classifier***

Local Context

- ***Token level*** encoder of the sentence to be simplified c_i

Global Context

- fixed window of Sentence level embedding (SBERT) for ***surrounding sentences***
- ***The left context is dynamically updated with previously simplified sentences***



Planning Model

RoBERTa classifier with cross-attention over the global context

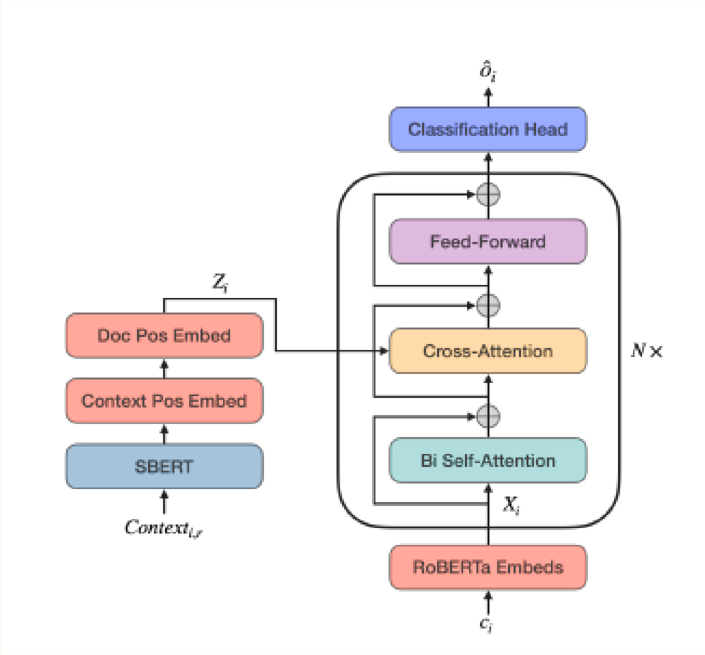
- *layers initialised with weights from a context-independent classifier*

Local Context

- ***Token level*** encoder of the sentence to be simplified c_i

Global Context

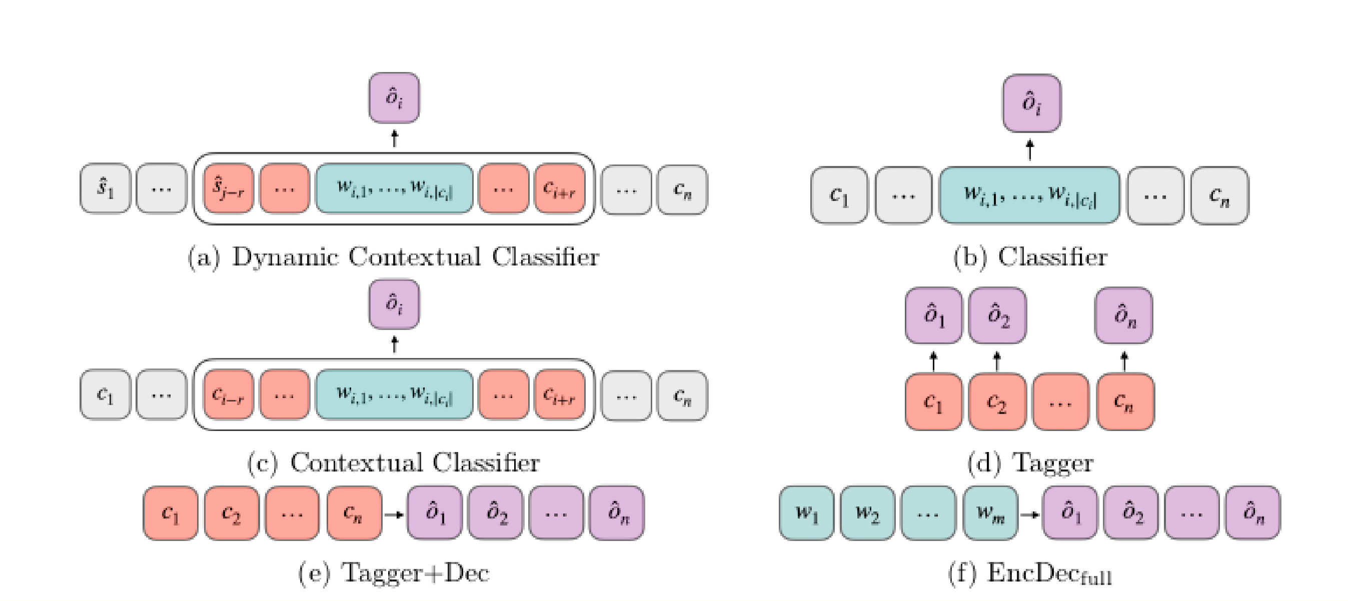
- fixed window of Sentence level embedding (SBERT) for ***surrounding sentences***
- *The left context is dynamically updated with previously simplified sentences*



Context positional embedding: relative distance of a given sentence from the input sentence c_i

Document positional embedding: the document quintile (1-5) that a given sentence falls into

Alternative Models



Dynamic Contextual Classifier: our model
Contextual Classifier: Static left context
Classifier: no context
Tagger: Sequence tagging on SBERT representations (no internal structure)

Tagger-Decoder: Each prediction is conditioned on the input document and on the previously predicted operation tags. SBERT encodings.
EncDec$_{full}$: Same as Tagger-Decoder but with token encodings

Data

(C, S) pairs with C a complex document and S its simplification.

Newsela

- News articles
- Each article is manually rewritten at five different levels of simplification, corresponding to discrete reading levels (0-4) of increasingly simplicity.
- Manual alignement of sentences and paragraphs

Wiki-auto

- Three simplification datasets which were automatically-collated from English Wikipedia and Wikipedia simple.
- Automatic alignement of sentences and paragraphs

Labeling the data

$(C, S) \rightarrow (C, S, o)$

Delete

- c_i is not aligned to any s_j .
  The complex sentence c_i is not aligned to any sentence s_j in the simplified version.

Labeling the data

$(C, S) \to (C, S, o)$

Delete

- c_i is not aligned to any s_j .

Copy

- c_i is aligned to a single s_j with a Levenshtein similarity above 0.92.
  The complex sentence c_i is aligned to a similar sentence s_j in the simplified version

Labeling the data

$(C, S) \rightarrow (C, S, o)$

Delete

- c_i is not aligned to any s_j .

Copy

- c_i is aligned to a single s_j with a Levenshtein similarity above 0.92.

Rephrase

- c_i is aligned to a single s_j with a Levenshtein similarity below 0.92.
  The complex sentence c_i is aligned to a sentence s_j in the simplified version but differs from it.

Labeling the data

$(C, S) \rightarrow (C, S, o)$

Delete

- c_i is not aligned to any s_j .

Copy

- c_i is aligned to a single s_j with a Levenshtein similarity above 0.92.

Rephrase

- c_i is aligned to a single s_j with a Levenshtein similarity below 0.92.

Split

- c_i is aligned to multiple s_j
  The complex sentence c_i is aligned to several sentences in the simplified version.

Data Filtering

Wiki-auto

- We clip all complex documents after the last aligned paragraph.
- We remove documents where more than 50% of aligned sentences are labelled as *delete*.

Wiki-auto and Newsela-auto

- We remove all articles that exceed 1024 tokens (so that we can fit them into a baseline BART generative model).

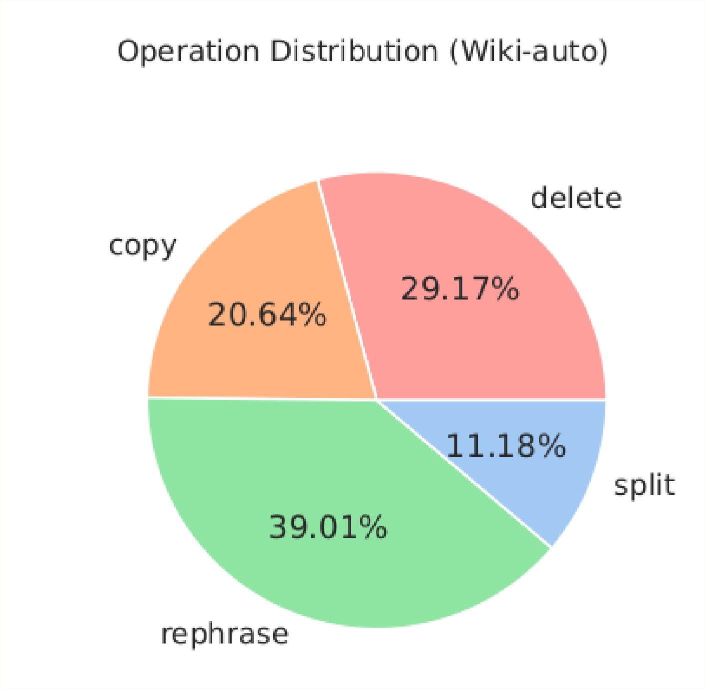
Data after filtering

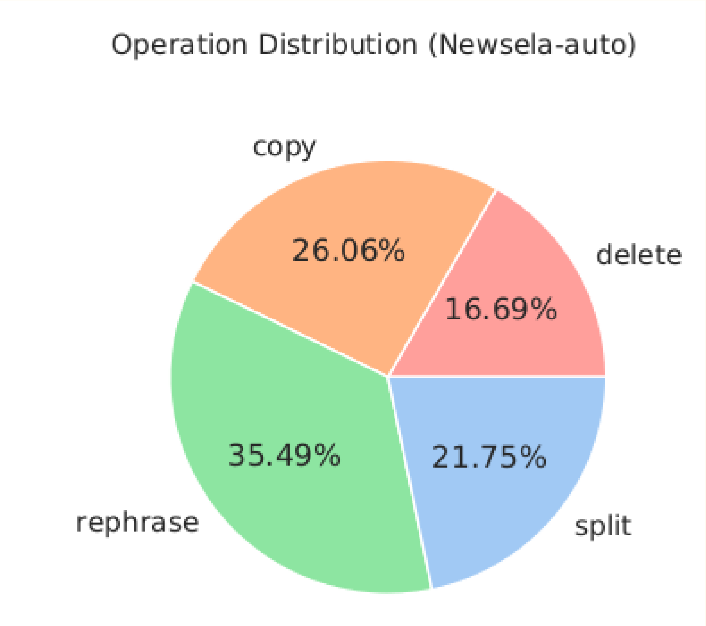
	Wiki-auto	Newsela-auto
# Doc Pairs	85,123	18,319
# Sent Pairs	461,852	707,776
Avg. $\|C\|$	155.51	868.98
Avg. $\|S\|$	97.72	674.94
Avg. $\|c_i\|$	28.64	22.49
Avg. $\|s_i\|$	21.57	15.84
Avg. n	5.43	38.64
Avg. k	4.53	42.60

- n: the number of sentences in C
- k: the number of sentences in S

- Newsela input documents are much longer
- Newsela data is smaller

Labelled Data





Planning Accuracy Results

	Wiki-auto						Newsela-auto					
Model	C	R	S	D	Micro	**Macro**	C	R	S	D	Micro	**Macro**
$\text{EncDec}_{\text{full}}$	26.9	42.2	36.0	51.8	43.2	40.8	26.1	10.8	11.7	9.0	12.2	11.5
EncDec	29.3	54.5	30.0	51.8	47.7	41.4	72.2	73.9	75.9	79.7	75.0	75.4
Tagger	38.6	54.2	31.7	**58.5**	50.6	45.8	71.4	72.7	74.1	78.4	73.7	74.1
Classifier	42.1	52.9	42.6	49.0	48.4	46.7	77.0	75.6	80.0	78.5	77.4	77.8
Dyn. Context	**44.8**	**57.9**	42.4	54.8	**52.8**	**50.0**	79.3	77.3	82.8	81.4	79.7	80.2
+ docpos	43.7	55.4	**43.6**	56.7	52.3	49.9	**80.0**	**78.1**	**83.6**	**82.0**	**80.3**	**80.8**

Our model consistently shows best results on both datasets.

Planning Accuracy Results

	Wiki-auto						Newsela-auto					
Model	C	R	S	D	Micro	**Macro**	C	R	S	D	Micro	**Macro**
$\text{EncDec}_{\text{full}}$	26.9	42.2	36.0	51.8	43.2	40.8	26.1	10.8	11.7	9.0	12.2	11.5
EncDec	29.3	54.5	30.0	51.8	47.7	41.4	72.2	73.9	75.9	79.7	75.0	75.4
Tagger	38.6	54.2	31.7	**58.5**	50.6	45.8	71.4	72.7	74.1	78.4	73.7	74.1
Classifier	42.1	52.9	42.6	49.0	48.4	46.7	77.0	75.6	80.0	78.5	77.4	77.8
Dyn. Context	**44.8**	**57.9**	42.4	54.8	**52.8**	**50.0**	79.3	77.3	82.8	81.4	79.7	80.2
+ docpos	43.7	55.4	**43.6**	56.7	52.3	49.9	**80.0**	**78.1**	**83.6**	**82.0**	**80.3**	**80.8**

The ***context-free classifier*** under-performs for ***Deletion***

- This confirms the intuition that ***global context*** particularly matters for that operation.

Planning Accuracy Results

	Wiki-auto						Newsela-auto					
Model	C	R	S	D	Micro	**Macro**	C	R	S	D	Micro	**Macro**
$\text{EncDec}_{\text{full}}$	26.9	42.2	36.0	51.8	43.2	40.8	26.1	10.8	11.7	9.0	12.2	11.5
EncDec	29.3	54.5	30.0	51.8	47.7	41.4	72.2	73.9	75.9	79.7	75.0	75.4
Tagger	38.6	54.2	31.7	**58.5**	50.6	45.8	71.4	72.7	74.1	78.4	73.7	74.1
Classifier	42.1	52.9	42.6	49.0	48.4	46.7	77.0	75.6	80.0	78.5	77.4	77.8
Dyn. Context	**44.8**	**57.9**	42.4	54.8	**52.8**	**50.0**	79.3	77.3	82.8	81.4	79.7	80.2
+ docpos	43.7	55.4	**43.6**	56.7	52.3	49.9	**80.0**	**78.1**	**83.6**	**82.0**	**80.3**	**80.8**

Sentence level encoding of the input sentence yields worse results (EncDec, Tagger)

- The loss is strongest for the ***Split*** operation
- This confirms the intuition that ***local context*** particularly matters for that operation.

Planning Accuracy Results

	Wiki-auto						Newsela-auto					
Model	C	R	S	D	Micro	**Macro**	C	R	S	D	Micro	**Macro**
EncDec$_{full}$	26.9	42.2	36.0	51.8	43.2	40.8	26.1	10.8	11.7	9.0	12.2	11.5
EncDec	29.3	54.5	30.0	51.8	47.7	41.4	72.2	73.9	75.9	79.7	75.0	75.4
Tagger	38.6	54.2	31.7	**58.5**	50.6	45.8	71.4	72.7	74.1	78.4	73.7	74.1
Classifier	42.1	52.9	42.6	49.0	48.4	46.7	77.0	75.6	80.0	78.5	77.4	77.8
Dyn. Context	**44.8**	**57.9**	42.4	54.8	**52.8**	**50.0**	79.3	77.3	82.8	81.4	79.7	80.2
+ docpos	43.7	55.4	**43.6**	56.7	52.3	49.9	**80.0**	**78.1**	**83.6**	**82.0**	**80.3**	**80.8**

A token level modeling of the document context performs worst (EncDecfull)

- This suggests that the very long input challenges the attention mechanism

Ablations

Model	Copy	Rephrase	Split	Delete	Micro	**Macro**
(a) Ablation on Best Model						
Dyn, $r = 13$, +init, +docpos	80.0	78.1	83.6	82.0	80.3	80.8
-docpos	79.3	77.3	82.8	81.4	79.7	80.2
-init	74.9	72.1	77.8	75.2	74.6	75.0
-init, -docpos	75.6	72.0	77.7	77.1	75.1	75.6
(b) Dynamic vs. Static Context						
Stat, $r = 9$	71.3	69.5	75.4	73.3	72.0	72.4
Stat, $r = 13$	72.2	65.3	69.9	68.3	68.5	68.9
Dyn, $r = 9$	73.1	70.1	75.5	75.9	73.1	73.6
Dyn, $r = 13$	75.6	72.0	77.7	77.1	75.1	75.6
(c) With vs without Initialisation						
Dyn, $r = 9$	73.1	70.1	75.5	75.9	73.1	73.6
Dyn, $r = 9$ +init	79.3	78.0	82.7	79.8	79.7	80.0
Dyn, $r = 13$	75.6	72.0	77.7	77.1	75.1	75.6
Dyn, $r = 13$ +init	79.3	77.3	82.8	81.4	79.7	80.2
(d) Window Size						
Stat, $r = 9$	71.3	69.5	75.4	73.3	72.0	72.4
Stat, $r = 13$	72.2	65.3	69.9	68.3	68.5	68.9
Dyn, $r = 9$	73.1	70.1	75.5	75.9	73.1	73.6
Dyn, $r = 13$	75.6	72.0	77.7	77.1	75.1	75.6
Dyn, $r = 9$ +docpos	73.8	72.9	77.2	75.8	74.6	74.9
Dyn, $r = 13$ +docpos	74.9	72.1	77.8	75.2	74.6	75.0
Dyn, $r = 9$ +init +docpos	79.4	78.0	83.1	82.0	80.1	80.6
Dyn, $r = 13$ +init +docpos	80.0	78.1	83.6	82.0	80.3	80.8

Plan-Guided Document Simplification

$$c_i, \hat{o}_i \Rightarrow s_i$$

Cripwell et al. EACL 2023

Plan Guided Document Simplification

Predict simplification operations

$$c1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$$

Simplify each input sentences using controls

$$c_i, \hat{o}_i \Rightarrow s_i$$

Document Simplification Models

BART Encoder-Decoder model fine-tuned on simplification data

Iterating over the document sentences

- Plan-Guided (PG): ***pipeline***

$$c_i, \hat{o}_i \Rightarrow s_i$$

Document Simplification Models

BART Encoder-Decoder model fine-tuned on simplification data

Iterating over the document sentences

- Plan-Guided (PG): ***pipeline***

  $$c_i, \hat{o}_i \Rightarrow s_i$$

- Sent-BART: ***end-to-end***

  $$c_i \Rightarrow s_i$$

Document Simplification Models

BART Encoder-Decoder model fine-tuned on simplification data

Iterating over the document sentences

- Plan-Guided (PG): ***pipeline***

$$c_i, \hat{o}_i \Rightarrow s_i$$

- Sent-BART: ***end-to-end***

$$c_i \Rightarrow s_i$$

- Doc-BART

$$DOC \Rightarrow SIMPLIFIED$$

Evaluation Metrics

SARI (Xu et al., 2016)

- Most popular simplification metric.
- Computes n-gram edits between input, output, and references.

Summarization metrics

- BARTScore (Yuan et al., 2021)
- SMART (Amplayo et al., 2022)

FKGL (Kincaid et al., 1975)

- Readibility metrics
- Uses surface-level statistics like syllable counts and sentence length.

Results

System	BARTScore ↑				SMART ↑			FKGL ↓	SARI ↑	Length	
	Faith. $(s \to h)$	P $(r \to h)$	R $(h \to r)$	F1	P	R	F1			Tokens	Sents
Input	-0.93	-2.47	-1.99	-2.23	63.2	62.7	62.8	8.44	20.52	866.9	38.6
Reference	-1.99	-0.93	-0.93	-0.93	100	100	100	4.93	99.99	671.5	42.6
Doc-BART	-2.48	-2.68	-2.76	-2.72	61.9	43.9	50.6	10.01	47.07	600.8	20.7
Sent-BART	**-1.86**	-1.63	-1.56	-1.60	78.9	80.1	79.3	5.03	73.02	666.4	42.6
PG_{Tag}	-1.95	-2.22	-2.18	-2.20	5.07	62.0	62.6	61.6	56.13	657.4	41.8
PG_{EncDec}	-1.94	-2.22	-2.18	-2.20	62.2	62.5	61.6	5.09	56.06	654.2	41.4
PG_{Clf}	-1.91	-1.68	**-1.53**	-1.60	77.8	**81.2**	79.3	**4.95**	73.83	688.8	44.5
PG_{Dyn}	-1.91	**-1.60**	-1.54	**-1.57**	**80.2**	81.0	**80.5**	4.98	**75.00**	667.2	42.6
PG_{Oracle}	-1.93	**-1.39**	**-1.40**	**-1.40**	**85.5**	**85.0**	**85.3**	**4.91**	**80.74**	655.6	42.1

- ***Our model*** (PG Dyn) achieves the highest results of all systems.

Results

System	BARTScore ↑				SMART ↑			FKGL ↓	SARI ↑	Length	
	Faith. $(s \rightarrow h)$	P $(r \rightarrow h)$	R $(h \rightarrow r)$	F1	P	R	F1			Tokens	Sents
Input	-0.93	-2.47	-1.99	-2.23	63.2	62.7	62.8	8.44	20.52	866.9	38.6
Reference	-1.99	-0.93	-0.93	-0.93	100	100	100	4.93	99.99	671.5	42.6
Doc-BART	-2.48	-2.68	-2.76	-2.72	61.9	43.9	50.6	10.01	47.07	600.8	20.7
Sent-BART	**-1.86**	-1.63	-1.56	-1.60	78.9	80.1	79.3	5.03	73.02	666.4	42.6
PG_{Tag}	-1.95	-2.22	-2.18	-2.20	5.07	62.0	62.6	61.6	56.13	657.4	41.8
PG_{EncDec}	-1.94	-2.22	-2.18	-2.20	62.2	62.5	61.6	5.09	56.06	654.2	41.4
PG_{Clf}	-1.91	-1.68	**-1.53**	-1.60	77.8	**81.2**	79.3	**4.95**	73.83	688.8	44.5
PG_{Dyn}	-1.91	**-1.60**	-1.54	**-1.57**	**80.2**	81.0	**80.5**	4.98	**75.00**	667.2	42.6
PG_{Oracle}	-1.93	**-1.39**	**-1.40**	**-1.40**	**85.5**	**85.0**	**85.3**	**4.91**	**80.74**	655.6	42.1

- ***Our model*** (PG Dyn) achieves the highest results of all systems.
- ***Improving planning*** (PG Oracle) would substantially increase performance (PG Oracle)

Results

System	BARTScore ↑				SMART ↑			FKGL ↓	SARI ↑	Length	
	Faith. ($s \rightarrow h$)	P ($r \rightarrow h$)	R ($h \rightarrow r$)	F1	P	R	F1			Tokens	Sents
Input	-0.93	-2.47	-1.99	-2.23	63.2	62.7	62.8	8.44	20.52	866.9	38.6
Reference	-1.99	-0.93	-0.93	-0.93	100	100	100	4.93	99.99	671.5	42.6
Doc-BART	-2.48	-2.68	-2.76	-2.72	61.9	43.9	50.6	10.01	47.07	600.8	20.7
Sent-BART	**-1.86**	-1.63	-1.56	-1.60	78.9	80.1	79.3	5.03	73.02	666.4	42.6
PG_{Tag}	-1.95	-2.22	-2.18	-2.20	5.07	62.0	62.6	61.6	56.13	657.4	41.8
PG_{EncDec}	-1.94	-2.22	-2.18	-2.20	62.2	62.5	61.6	5.09	56.06	654.2	41.4
PG_{Clf}	-1.91	-1.68	**-1.53**	-1.60	77.8	**81.2**	79.3	**4.95**	73.83	688.8	44.5
PG_{Dyn}	-1.91	**-1.60**	-1.54	**-1.57**	**80.2**	81.0	**80.5**	4.98	**75.00**	667.2	42.6
PG_{Oracle}	-1.93	**-1.39**	**-1.40**	**-1.40**	**85.5**	**85.0**	**85.3**	**4.91**	**80.74**	655.6	42.1

- ***Our model*** (PG Dyn) achieves the highest results of all systems.
- ***Improving planning*** (PG Oracle) would substantially increase performance (PG Oracle)
- End-to-end ***simplification of the full document*** (Doc-BART) yields poor results

Example output

Complex	Silvano "Nano" Campeggi (January 23, 1923 – August 29, 2018) was an Italian artist who designed and produced the artwork for the posters of many classic Hollywood films. His iconic images are associated with the golden era of Hollywood and Campeggi is now generally regarded as the most important graphic artist and poster designer in the history of American cinema. In the following decades, Campeggi designed and produced the poster and advertising graphics for over 3000 films, working not only under contract with the MGM studios, but also with Warner Brothers, Paramount, Universal, Columbia Pictures, United Artists, RKO, Twentieth-Century Fox and several other movie studios. Sixty-four of the films he illustrated won Oscars, including "Casablanca", "Ben-Hur", "Singin' in the Rain", "An American in Paris", "West Side Story", "Exodus", "Breakfast at Tiffany's", and "Gigi". Campeggi died on 29 August 2018, at the age of 95.
Simple	Silvano "Nano" Campeggi (January 23, 1923 – August 29, 2018) was an Italian artist. He designed and produced the artwork for the posters of many classic Hollywood movies. Campeggi was known for his poster designs for "Casablanca", "Singin' in the Rain", and "Breakfast at Tiffany's". Campeggi died on August 29, 2018 in Florence from respiratory failure at the age of 95.
Doc-BART	Silvano "Nano" Campeggi (January 23, 1923 – August 29, 2018) was an Italian artist. He designed and produced the artwork for the posters of many classic Hollywood movies, including "Casablanca", "Ben-Hur", "Singin' in the Rain", "An American in Paris", "West Side Story", "Exodus", "Breakfast at Tiffany's", "Gigi", and "".
Sent-BART	Silvano "Nano" Campeggi (23 January 1923 – 29 August 2018) was an Italian artist. He designed and produced the artwork for the posters of many classic Hollywood movies. Campeggi is generally regarded as the most important graphic artist and poster designer in the history of American cinema. Campeggi designed and produced the poster and advertising graphics for over 3000 movies, working not only under contract with the MGM studios, but also with Warner Brothers, Paramount, Universal, Columbia Pictures, United Artists, RKO, Twentieth Century Fox and several other movie studios. Sixty-four of the movies he illustrated won Oscars, including "Casablanca", "Ben-Hur", "Singin' in the Rain", "An American in Paris", "West Side Story", "Exodus", "Breakfast at Tiffany's", and "Gigi". Campeggi died on 29 August 2018, at the age of 95.
PG_{Dyn}	Silvano "Nano" Campeggi (January 23, 1923 – August 29, 2018) was an Italian artist. He designed and produced the artwork for the posters of many classic Hollywood movies. Sixty-four of the movies he illustrated won Oscars, including "Casablanca", "Ben-Hur", "Singin' in the Rain", "An American in Paris", "West Side Story", "Exodus", "Breakfast at Tiffany's", and "Gigi". Campeggi died on 29 August 2018 at the age of 95.

Context-Aware Document Simplification

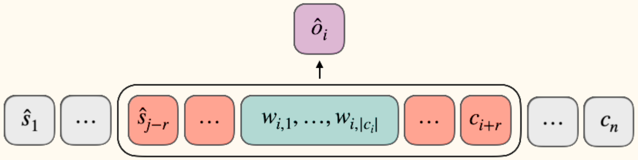
Cripwell et al. Findings of ACL 2023

Context-Aware Simplification

PG (plan-guided) pipeline

First PLAN,
Input D $\Rightarrow$ Simplification Plan
$c1, \ldots, c_n \Rightarrow \hat{o}, \ldots, \hat{o}_n$

PLANNING is Context-Aware ...

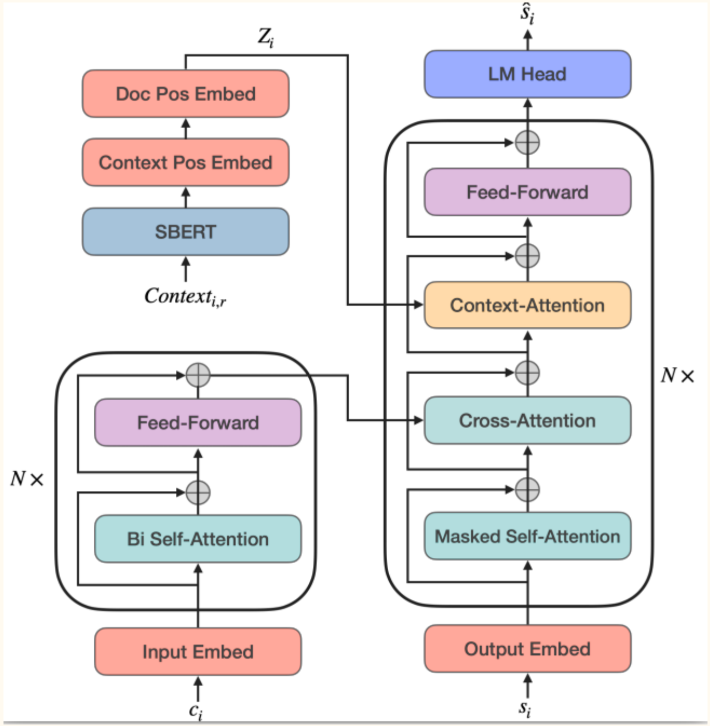


then SIMPLIFY
Input S + Simplification Operation $\Rightarrow$ Simplified S
$c_i, \hat{o}_i \Rightarrow s_i$

... but SIMPLIFICATION is not

Context-Aware BART (ConBART)

- Modification of the BART architecture
- Generation is ***conditioned on both an input sentence c_i and a representation of the document context Z_i*** of that sentence
- Same ***context modeling*** as for planner (SBERT encoding of the neighbouring sentences)



Contexts and Models

Textual inputs at varying granularities

- BARTdoc, BARTpara, BARTsent, LEDdoc, LEDpara

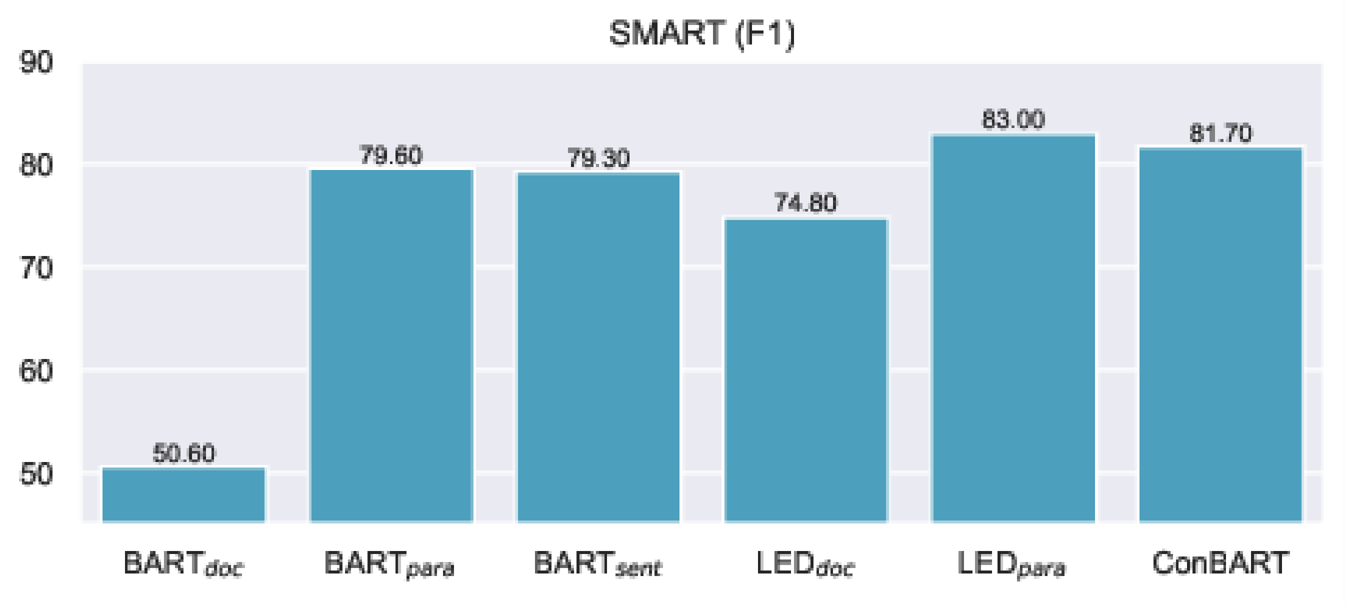
Complex sentence input + Global Context

- ConBART

All above systems + plan-guidance ($\hat{O} \rightarrow M$)

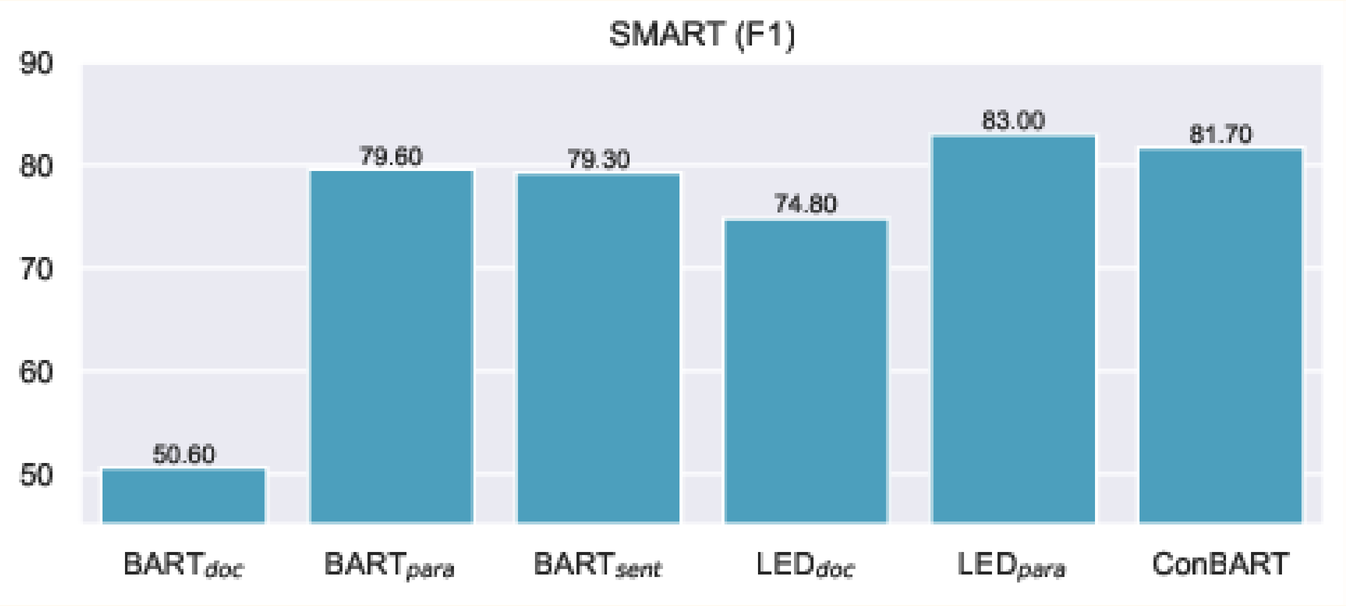
- $\hat{O}$, a predicted simplification plan
- M, a simplification model (BART, LED, ConBART)

Which context helps most ?



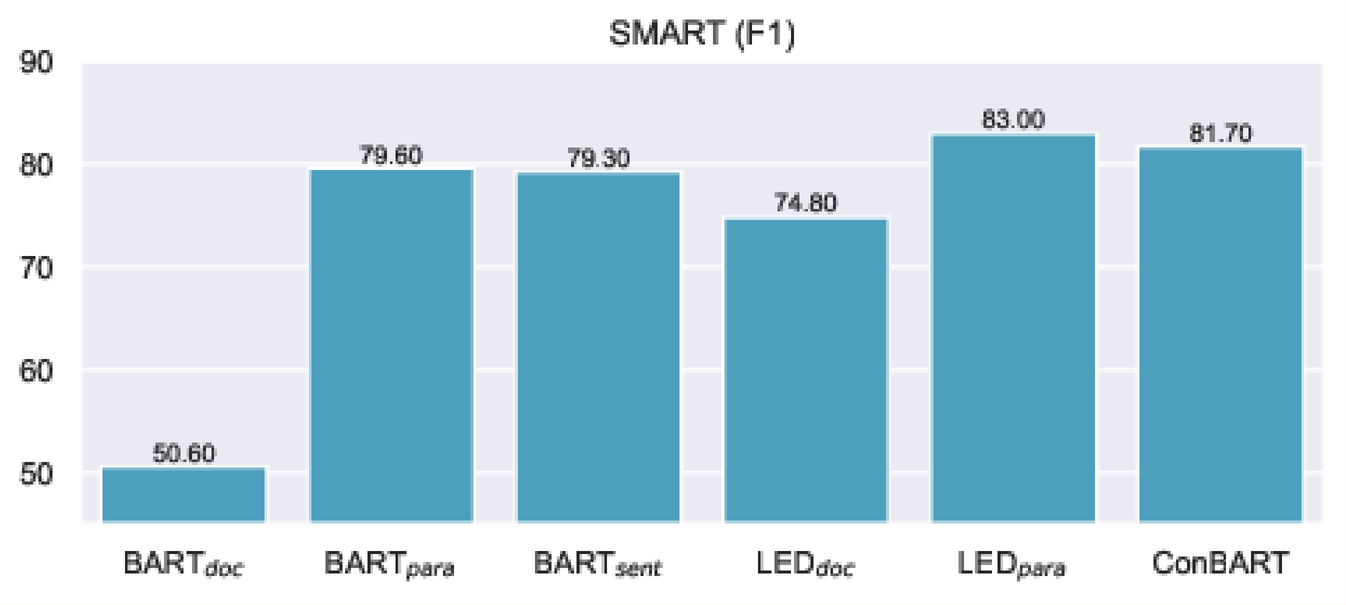
The best two models use a medium size context (either a paragraph or a sentence window)

Which context helps most ?



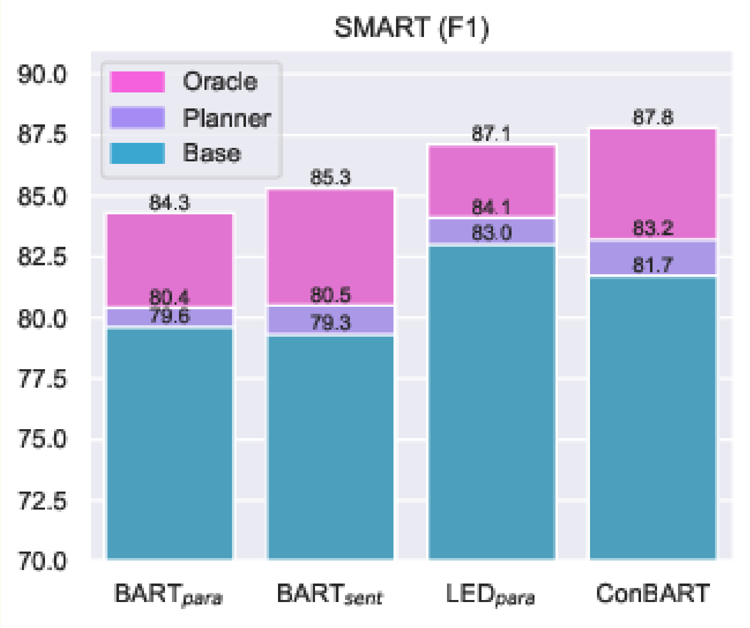
Full Document context does not work well (BARTdoc, LEDdoc)

Which context helps most ?



For end-to-end models, LongFormers drastically improve results on longer input (document, paragraph)

Does planning help ?

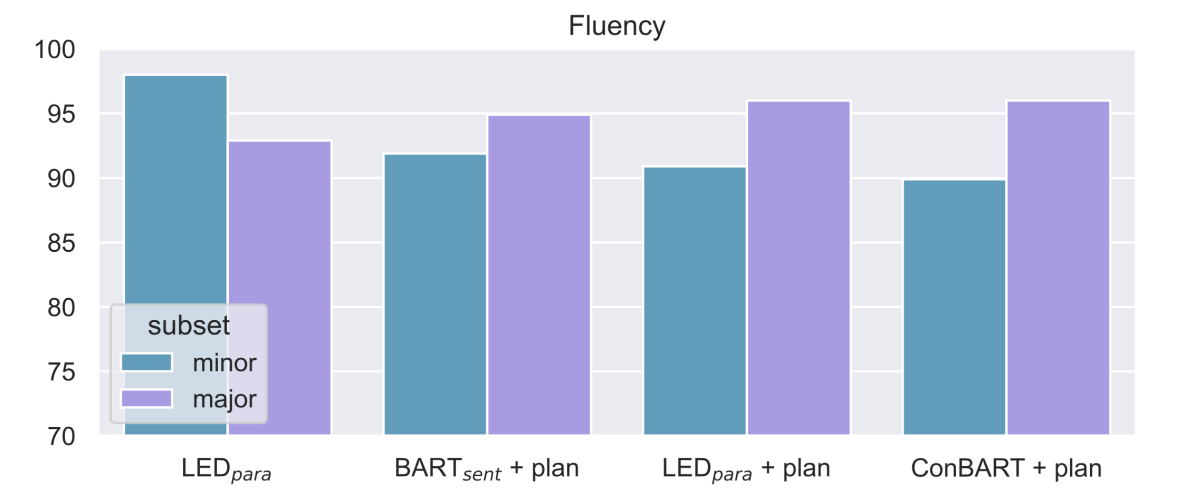


- Planning systematically improves performance
- Planning needs improving
  - the model simplifying based on the oracle plan has much higher performance

Human Evaluation

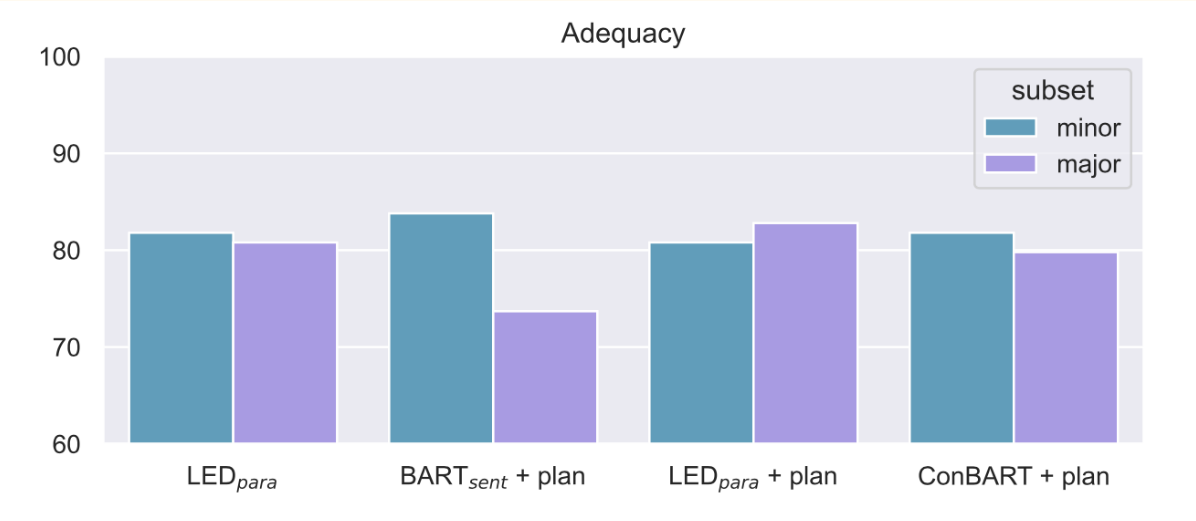
- On paragraphs
  - 33 complex paragraphs from each non-adjacent reading-level transition pairing
  - 198 paragraphs in total
  - 50% Minor: reading-level transition of two (0-2, 1-3 etc)
  - 50% Major: reading-level transition higher than two (0-3, 1-4 etc)
- Yes/No judgments on fluency, adequacy, simplicity
- Score = proportion of positive judgments
- References and outputs from 4 high performing systems
  - PGDyn, LEDpara, $\hat{O} \rightarrow LEDpara, \hat{O} \rightarrow ConBART$)
- 990 outputs in total

Human Evaluation



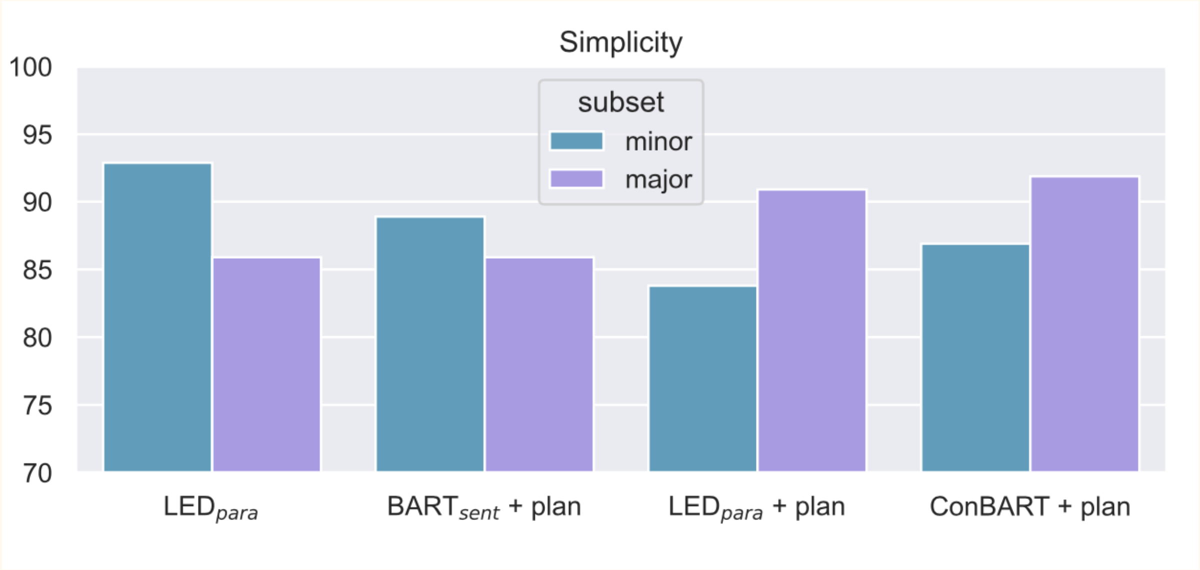
- All systems achieve high fluency – not surprising given modern LM
- Planning improves fluency on MAJOR cases (cases requiring higher degrees of simplification)

Human Evaluation



- Window- (ConBART) and paragraph-based models are better at maintaining adequacy.

Human Evaluation



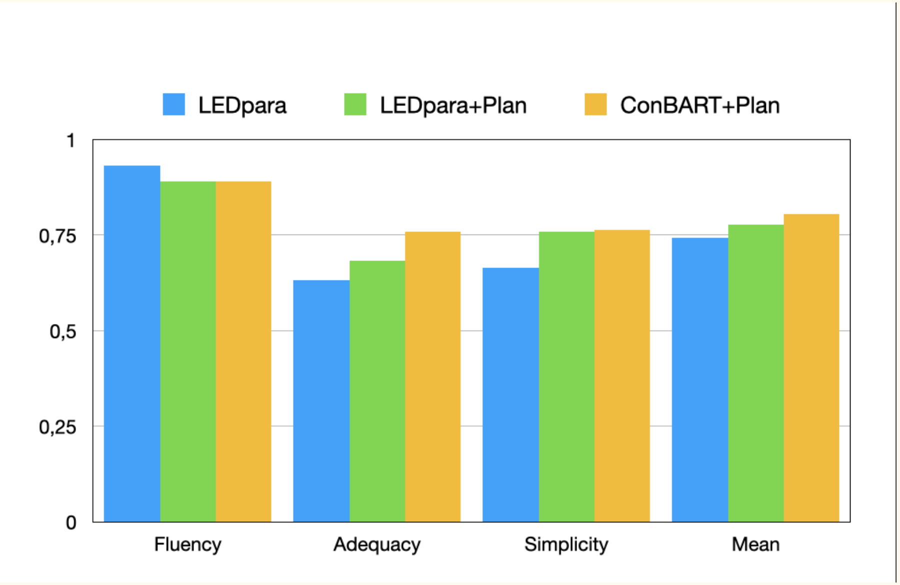
- Window/paragraph-based models + Planning yields high simplicity in major cases (overcoming conservativity?)
    - (LEDpara/ConBART + plan)

Generalising to OOD Data

Training on Newsela, Testing on Wiki-auto

Generalising to OOD Data

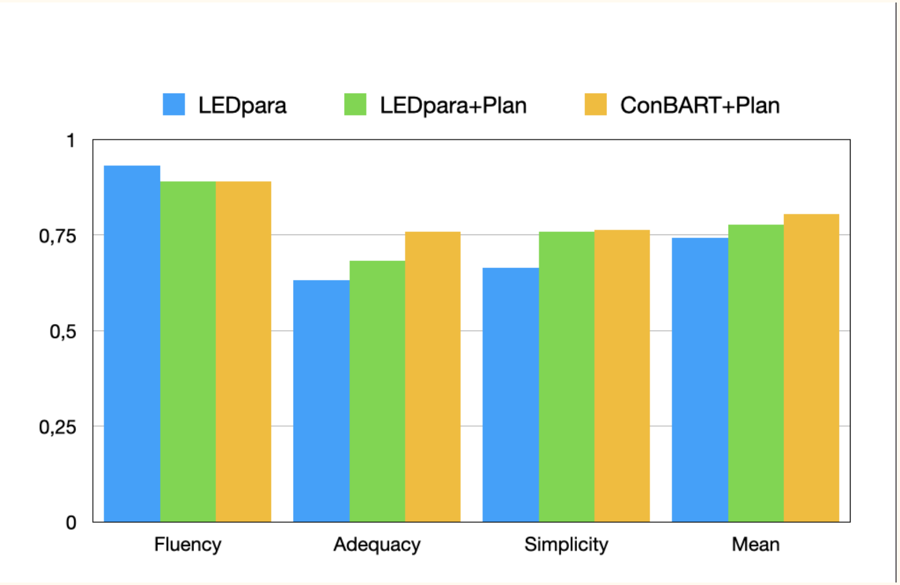
Training on Newsela, Testing on Wiki-auto



- Planning helps on unseen domains.

Generalising to OOD Data

Training on Newsela, Testing on Wiki-auto



- Planning helps on unseen domains.
- Paragraph-based models are less adaptable to unseen domains
  - Paragraph length varies across corpora making
  - Models tend to be biased towards paragraph length of training data

Conclusion and Perspectives

Conclusion and Future Work

Planning Simplification operations and having a window-based context helps

- improve document simplification
- generalising to new domains
- handling more drastic simplification (MAJOR cases)

Simplification metrics

- there is a need for a reference less metric which correctly captures the tradeoff between meaning preservation and simplification

Types of Simplification

- Here (Newsela): simplification in terms of school level
- What about: expert/layman, disadvantaged users ?

LLMs

- How well do they simplify ?
- Can prompting helps diversifying simplification (generate simplifications for diverse users)?

Questions ?

Example Output

Input

He was born in Stavanger; his father was a military engineer and he was the grandson of the historian Christian C. A. Lange. He graduated from *secondary school* in 1887 and proceeded to travel and study history, English, and French at the University of Oslo, from which he received the cand.philol. degree in 1893. He taught at secondary schools for many years and eventually returned to the University of Oslo to receive a doctorate.

Output

Almond was born in Stavanger, Norway. <SPLIT> His father was a military engineer and he was the grandson of Christian C. A. Lange. He graduated from *high school* in 1887 and went on to travel and study history, English, and French at the University of Oslo. <SPLIT> In 1893 he received a doctorate in physics. He taught at secondary schools for many years and eventually returned to the University of Oslo to receive a doctorate.

Example Output

Input

Historical research indicates that the "Zibelemärit" *originated in the 1850s* with "marmettes", farmer's wives from around Murten, coming to Bern at around St. Martin's Day to sell their produce; however, a persistent local legend *holds* that the "Zibelemärit" is a much older *festivity*. According to this legend, the Bernese *awarded* the people from the nearby city of Fribourg the right to sell onions in the city in reward for their aid after a fire destroyed much of Bern in 1405.

Output

The "Zibelemärit" *started around 150 years ago* with “marmettes", farmer's wives. <SPLIT> They came to Bern at around St. Martin's Day to sell their produce. <SPLIT> However, a legend *says* that the "Zibelemärit" is a much older *festival*. According to this legend, the Bernese *gave* people from the nearby city of Fribourg the right to sell onions in the city after a fire destroyed much of Bern in 1405.

Example Output

Input

The Zibelemärit is an annual market with aspects of a fair in the old town of Bern, Switzerland. It takes place the fourth Monday in November.

Historical research indicates that the "Zibelemärit" *originated in the 1850s* with "marmettes", farmer's wives from around Murten, coming to Bern at around St. Martin's Day to sell their produce; however, a persistent local legend *holds* that the "Zibelemärit" is a much older *festivity*. According to this legend, the Bernese *awarded* the people from the nearby city of Fribourg the right to sell onions in the city in reward for their aid after a fire destroyed much of Bern in 1405.

As the name indicates, it is mainly onions that are sold on the "Zibelemärit". Bernese farmers, who are proud of their decorative onion tresses and onion wreaths, also sell other onion products on the market, including Zwiebelkuchen (onion pie), onion soup and onion sausages. Decorative chains of sugar onions are also popular with children.

The "Zibelemärit" opens very early in the day, at around 03:00 to 04:00. Later in the morning, the narrow alleys are usually packed tight with people, which is what the Bernese call the "Gstungg". A general confetti battle in which mostly children *participate ensues* at four o'clock in the afternoon, *officially ending the market*.

Output

The Zibelemärit is an annual market in the old town of Bern, Switzerland. It takes place the fourth Monday in November.

The "Zibelemärit" *started around 150 years ago* with "marmettes", farmer's wives. <SPLIT> They came to Bern at around St. Martin's Day to sell their produce. <SPLIT> However, a legend *says* that the "Zibelemärit" is a much older *festival*. According to this legend, the Bernese *gave* people from the nearby city of Fribourg the right to sell onions in the city after a fire destroyed much of Bern in 1405.

In this country, it is mainly onions that are sold on the "Zibelemärit." Bernese farmers also sell other products, including Zwiebelkuchen (onion pie), onion soup and onion sausages. Decorative chains of sugar onions are also popular with children.

The "Zibelemärit" opens very early in the day, at around 03:00 to 04:00. Later in the morning, the narrow alleys are usually packed tight with people, which is what the Bernese call the "Gstungg." A general confetti battle in which mostly children *fight breaks out* at four o'clock in the afternoon *to end the market*.

Example Output

Input

Glenn Edward Greenwald (born March 6, 1967) is an American journalist and author.

He is best known for a series of reports published from June 2013 by "The Guardian" newspaper detailing the United States and British global surveillance programs, and based on *classified documents disclosed* by Edward Snowden. Greenwald and the team he worked with won both a George Polk Award and a Pulitzer Prize for those reports.

He has written several best-selling books, including "No Place to Hide". Before the Snowden file *disclosures*, Greenwald was *considered one of the most influential* opinion columnists in the United States. After working as a constitutional attorney for ten years, he began *blogging* on national security issues before becoming a "Salon" *contributor* in 2007 and *then* for "The Guardian" in 2012. He now writes for (and has co-edited) "The Intercept", which he founded in 2013 with Laura Poitras and Jeremy Scahill.

Greenwald's work on the Snowden story was featured in the documentary "Citizenfour", which won *the 2014 Academy Award for Best Documentary Feature*. Greenwald *appeared on-stage with director Laura Poitras and Snowden's girlfriend, Lindsay Mills, when the Oscar was given*. In the 2016 Oliver Stone feature film "Snowden", Greenwald was played by actor Zachary Quinto.

Output

Glenn Greenwald is an American journalist and author.

He is best known for a series of reports published from June 2013 by the Guardian newspaper. <SPLIT> They are based on *documents leaked* by Edward Snowden.

He has written several best-selling books, including "No Place to Hide." Before the Snowden file *leaks*, Greenwald *was one of the most respected* opinion columnists in the United States. He began writing about national security issues before becoming a "Salon" *writer* in 2007 and *a writer* for "The Guardian" in 2012. He now writes for The Guardian.

Greenwald's work on the Snowden story was featured in the documentary "Citizenfour". <SPLIT> The movie won *an Academy Award*. Greenwald *worked with director Laura Poitras and Snowden's girlfriend, Lindsay Mills, to make the documentary*. The 2016 Oliver Stone feature, "Snowden," was played by Zachary Quinto.