

PAIR: Planning and Iterative Refinement in Pre-trained Transformers for Long Text Generation

Xinyu Hua¹ and Lu Wang²

¹Northeastern University

²University of Michigan



Northeastern University



Motivation

- Large pre-trained Transformers have shown great promise for text generation applications.

Motivation

If you could have dinner with one person, dead or alive, who would that be?

That's a tough one. I think I'd have to go with Steve Jobs. I'd love to pick his brain.

Ah, interesting choice. What do you know about Steve Jobs?

He was one of the most influential people in the history of technology. He was a visionary.

What questions would you ask him?

I'd want to know how he came up with the idea for the iPod. It was so revolutionary at the time.

Totally...the iPod changed the way everyone listens to music.

It really did. It's crazy to think that it was released in 2001. That was a long time ago.

Omg wow, that makes me feel old!

I know, right? Time flies by so fast. I wish I could go back to those simpler times.

Me too, friend, me too. Speaking of music, what's your favorite album to listen to lately?

I've been listening to a lot of Led Zeppelin lately. They're my favorite band.

[Roller+, 2020]

ran
her

News You Can't Use

All the AI-Generated Fake News That's Fit To Print

October 25, 2020

Breaking News Presidential Election of 2020 Murders, Attempted Murders and Homicides Law and Legislation Politics and Government

HOT TOPICS

- [Trump, Donald J](#)
- [United States Politics and Government](#)
- [Breaking News](#)
- [Investigations](#)
- [Criminal Profiles](#)
- [Presidential Election of 2020](#)
- [Biden, Joseph R Jr](#)
- [Impeachment](#)
- [Democratic Party](#)
- [Ukraine](#)


LATEST HEADLINES

High-end real estate agents allegedly charge a \$2,850 fee for apartments they target for low-income buyers

Manny Castro's mother bought him an apartment in a new rental complex in northwest Miami. The first time she had spotted one in the area,...

Posts tagged as "Presidential Election of 2020"

Biden's Motto Was 'If Not Now, When?'



Joe Biden's announcement that he is running for president on Friday came as no surprise to his Democratic rivals in Iowa.

"You can imagine this, Biden came here before, and we had all the math worked out that we could win this one," said Sen. Cory Booker, who is campaigning this weekend in the state.

ABOUT

Everything you are reading is machine-generated and not true!

All the news on this website is generated by AI as a demonstration of how fake news can be mass produced by computers.

Every word on this site is 100% computer generated (except for this text block). There is absolutely no human editing.

Think it is obvious that these stories are fake? [Try the quiz!](#)

[To learn how this works and get the code to generate your own fake news site, read the article.](#)

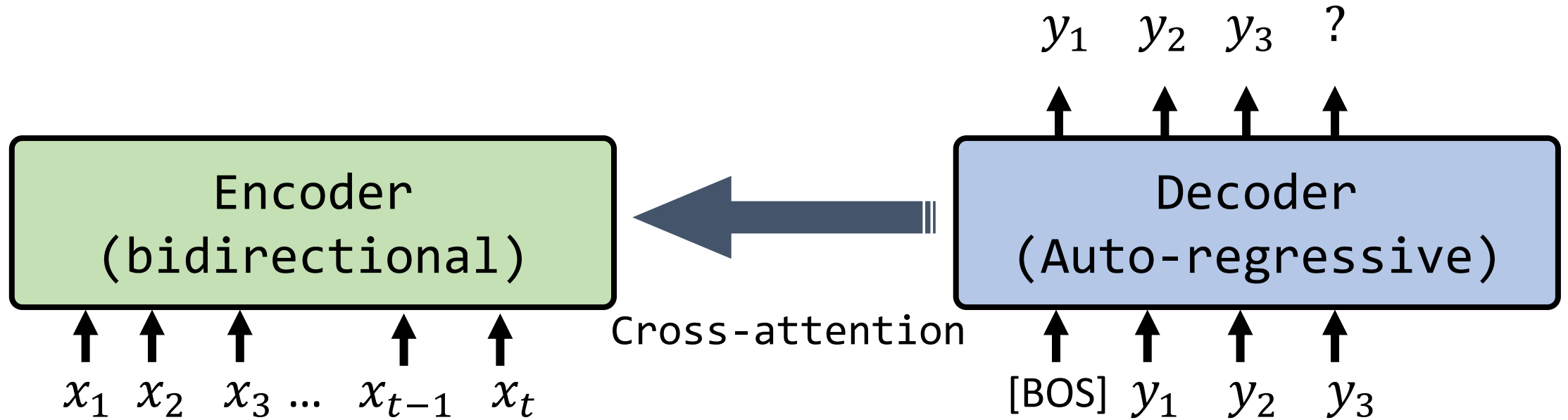
SPECIAL INVESTIGATION INTO JOHN MCFAKESON

Man pleads guilty to illegal killing of 60 kangaroos in his care

<https://newsyoucantuse.com/>

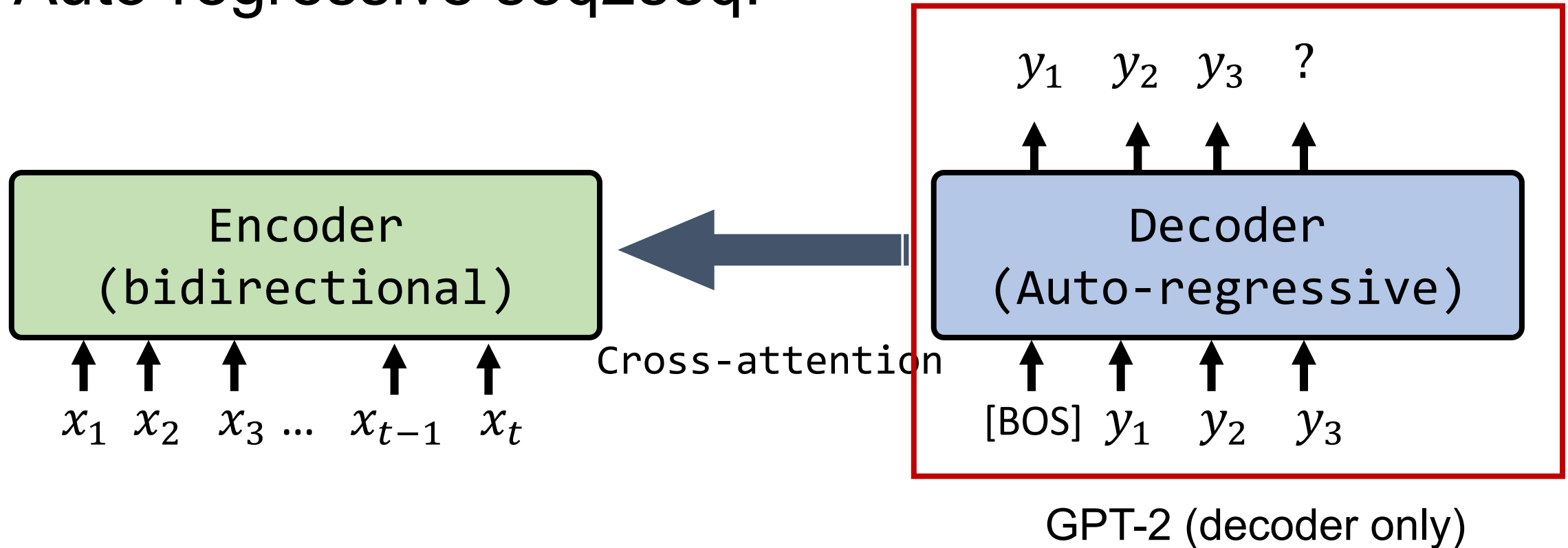
Motivation

- Auto-regressive seq2seq:



Motivation

- Auto-regressive seq2seq:

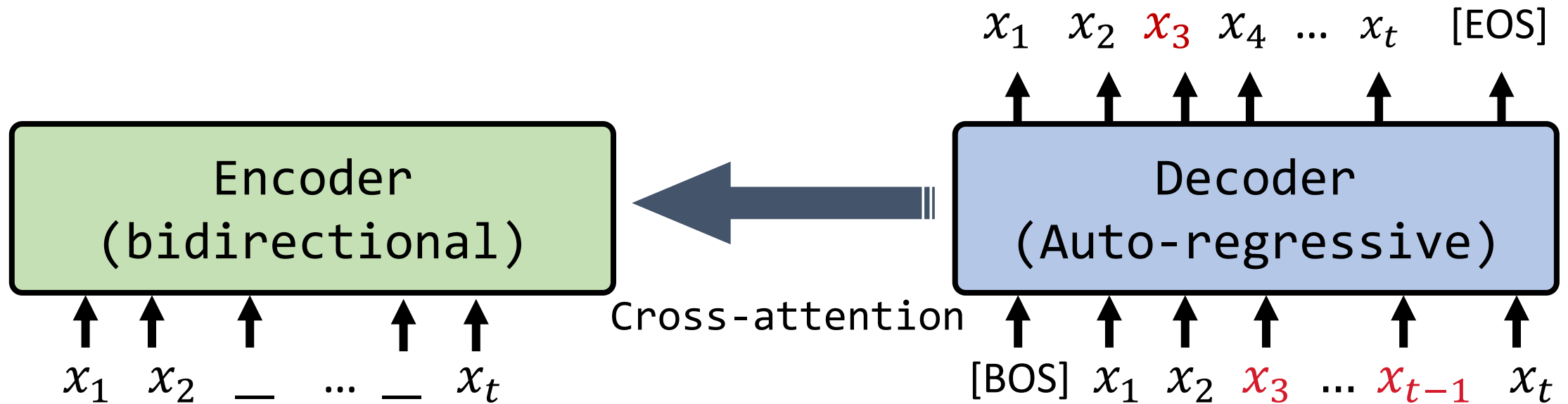


Major Challenges

- **Lack of coherent text plans** [[Wiseman+, 2018](#); [Moryossef+, 2019](#); [Hua & Wang, 2019](#)]
- **Low controllability** [[Wen+, 2015](#); [Hu+, 2017](#); [Dathathri+, 2020](#); [Goyal & Durrett, 2020](#)]
- **Hard to correct errors** [[Guu+, 2018](#); [Lee+, 2018](#); [Ghazvininejad+, 2019](#)]

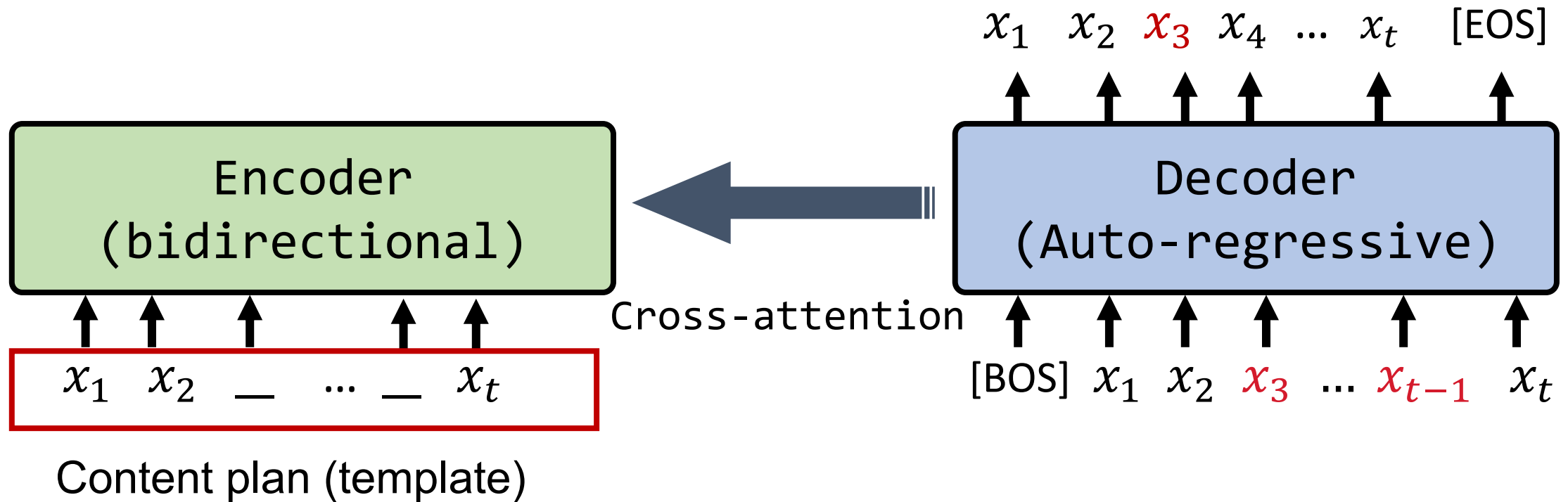
Motivation

- Masked language model:



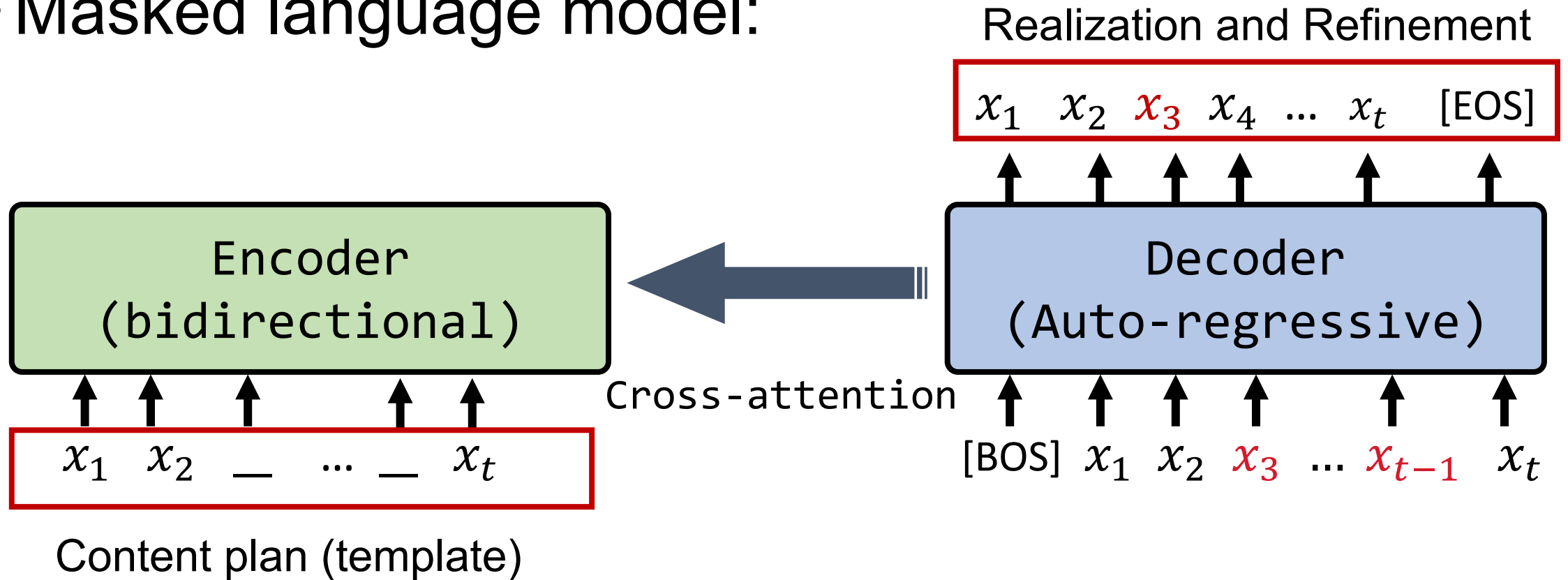
Motivation

- Masked language model:



Motivation

- Masked language model:



Proposal

- Prompt: *Donald Trump is a communist.*

- Content Plan:

S1: *A communist* ▶ *begin with* ▶ *coherent ideology* ▶ [SEN]

S2: [SEN]

S3: *No evidence* ▶ *any coherent* ▶ *held beliefs* ▶ *any topic* ▶ [SEN]

- Template:

S1: ___ ___ ___ *a communist* ___ ___ ___ *begin with* ___ ___ ___ ___ *coherent ideology* ___

___ ___ ___

S2: ___ ___ ___ ___

S3: ___ ___ *no evidence* ___ ___ ___ ___ *any coherent* ___ ___ *held beliefs* ___ *any topic* ___

Proposal

- **Template:**

S1: _____ a communist _____ begin with _____ coherent ideology _____

S2: _____

S3: _____ no evidence _____ any coherent _____ held beliefs _____ any topic _____

- **Draft: (mask-and-fill)**

S1: well call him a communist you must begin with that Donald Trump has coherent ideology to begin with.

S2: Which is unlikely.

S3: There is no evidence to suggest Donald Trump has any coherent or commonly held beliefs on any topic .

Proposal

- **Draft:** (mask-and-fill)

S1: we'll call him **a communist** you must **begin with** that Donald Trump has **coherent ideology** to begin with.

S2: Which is unlikely.

S3: There is **no evidence** to suggest Donald Trump has **any coherent** or commonly **held beliefs** on **any topic** .

- **Refined:** (Iterative Refinement)

S1: To call him **a communist** you must **begin with** that he has some kind of **coherent ideology** in the first place.

S2: He does not.

S3: There is **no evidence** whatsoever that Trump has **any coherent** or commonly **held beliefs** on **any topic** .

Roadmap

- Motivation
- **Related Work**
- PAIR model
- Experiments
- Results
- Conclusion

Related Work

- **Text Planning in Neural Text Generation**
 - Plan as sequences or trees [[Moryossef+, 2019](#); [Puduppully+, 2019](#); [Hua & Wang, 2019](#)]
- **Controlled Text Generation**
 - Topic, style, entities [[Wang+, 2017](#); [See+, 2019](#); [Fan+, 2018](#)]
 - Syntax and semantics [[Dusek & Jurcicek, 2016](#); [Wen+, 2015](#)]
- **Iterative Refinement**
 - Non-autoregressive machine translation [[Lee+, 2018](#); [Freitag+, 2019](#); [Mansimov+, 2019](#); [Kasai+, 2020](#)]
 - Masked language model [[Lawrence+, 2019](#); [Ghazvininejad+, 2019](#)]

Related Work

- **Text Planning in Neural Text Generation**
 - Plan as sequences or trees [Moryossef+, 2019; Puduppully+, 2019; Hua & Wang, 2019]
- **Controlled Text Generation**
 - Topic, style, entities [[Wang+, 2017](#); [See+, 2019](#); [Fan+, 2018](#)]
 - Syntax and semantics [[Dusek & Jurcicek, 2016](#); [Wen+, 2015](#)]
- **Iterative Refinement**
 - Non-autoregressive machine translation [Lee+, 2018; Freitag+, 2019; Mansimov+, 2019; Kasai+, 2020]
 - Masked language model [Lawrence+, 2019; Ghazvininejad+, 2019]

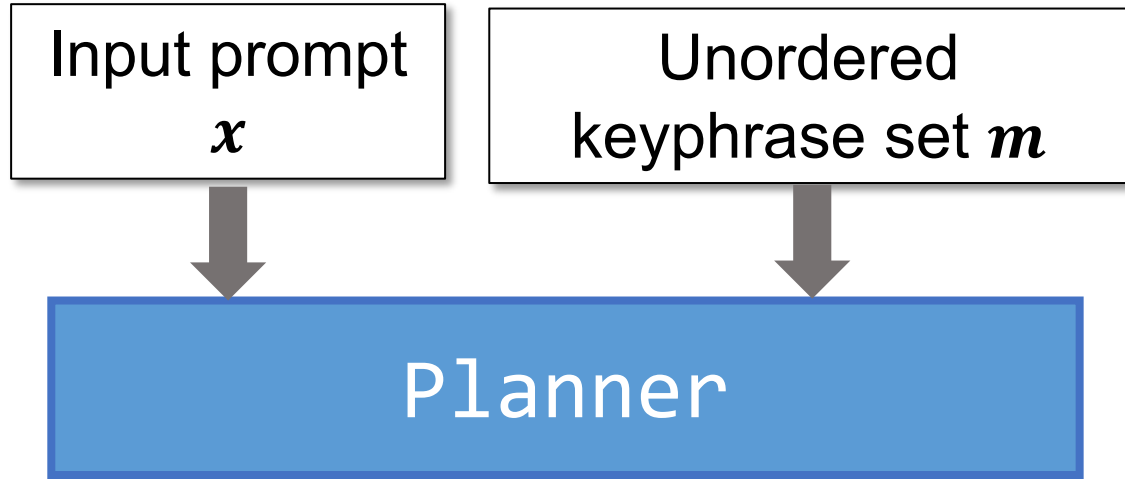
Related Work

- **Text Planning in Neural Text Generation**
 - Plan as sequences or trees [Moryossef+, 2019; Puduppully+, 2019; Hua & Wang, 2019]
- **Controlled Text Generation**
 - Topic, style, entities [Wang+, 2017; See+, 2019; Fan+, 2018]
 - Syntax and semantics [Dusek & Jurcicek, 2016; Wen+, 2015]
- **Iterative Refinement**
 - Non-autoregressive machine translation [Lee+, 2018; Freitag+, 2019; Mansimov+, 2019; Kasai+, 2020]
 - Masked language model [Lawrence+, 2019; Ghazvininejad+, 2019]

Roadmap

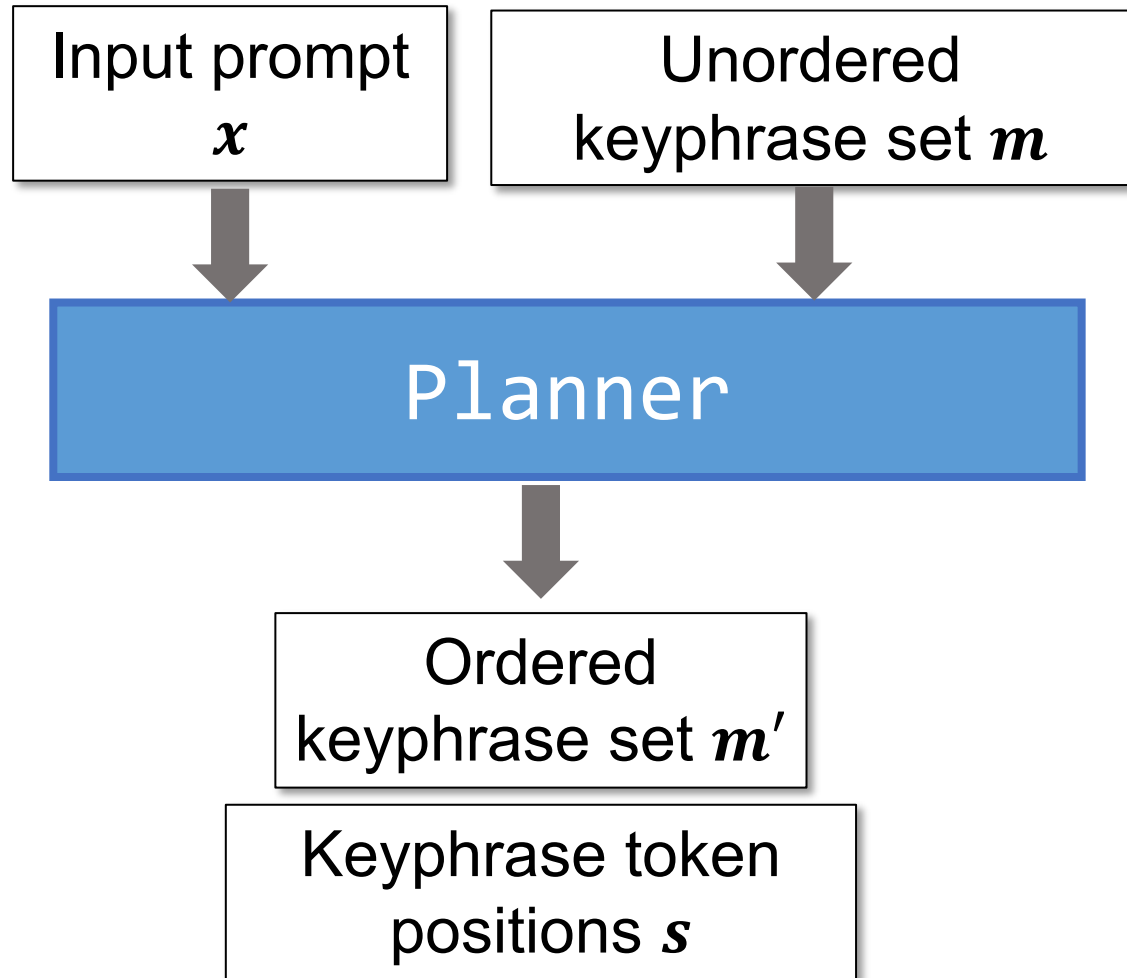
- Motivation
- Related Work
- **PAIR model**
- Experiments
- Results
- Conclusion

PAIR Model



Prompt: *CMV. Donald Trump is a communist.*
Keyphrase set: *{any coherent; any topic; begin with; coherent ideology; held beliefs ...}*

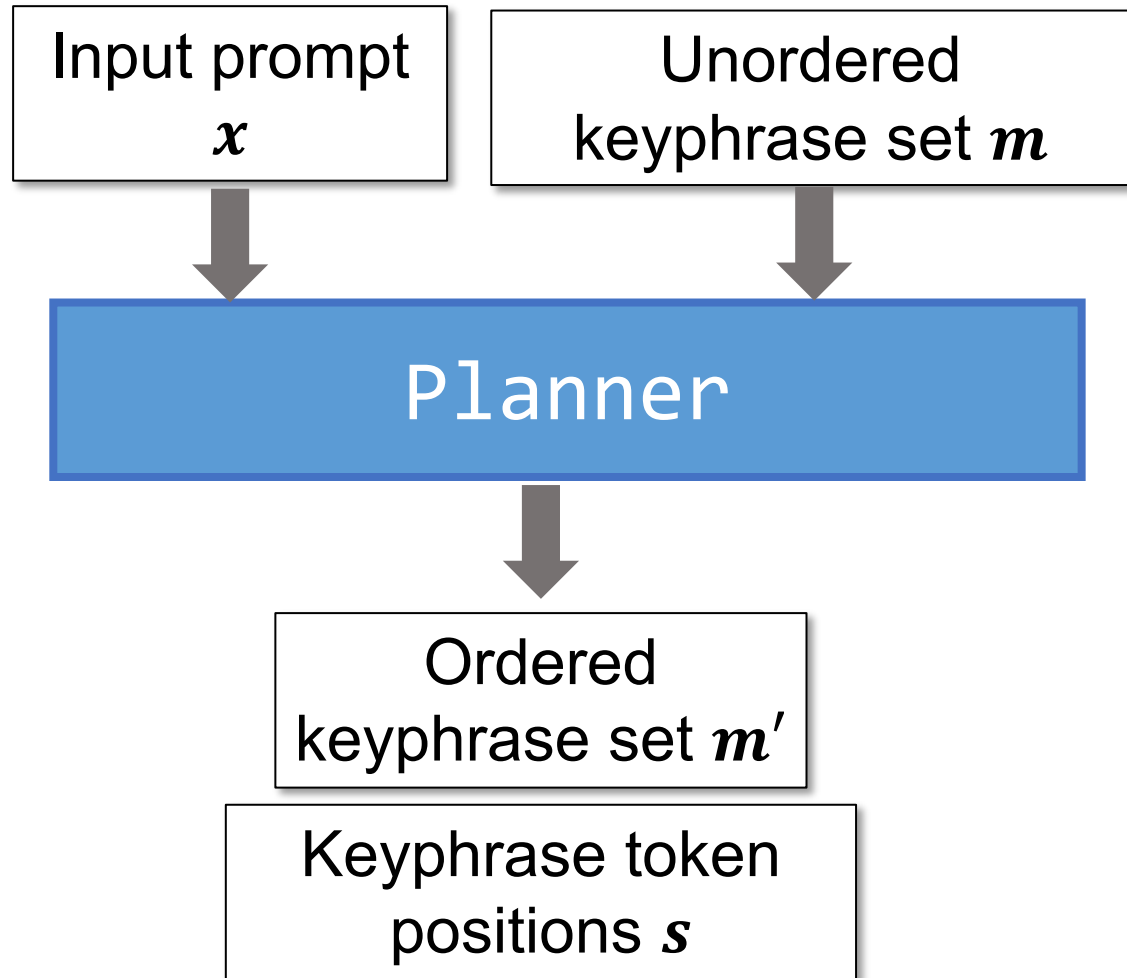
PAIR Model



Prompt: *CMV. Donald Trump is a communist.*
Keyphrase set: {*any coherent; any topic; begin with; coherent ideology; held beliefs ...*}

Text Plan: (a communist)₃ (begin with)₈ (coherent ideology)₁₅ ...

PAIR Model

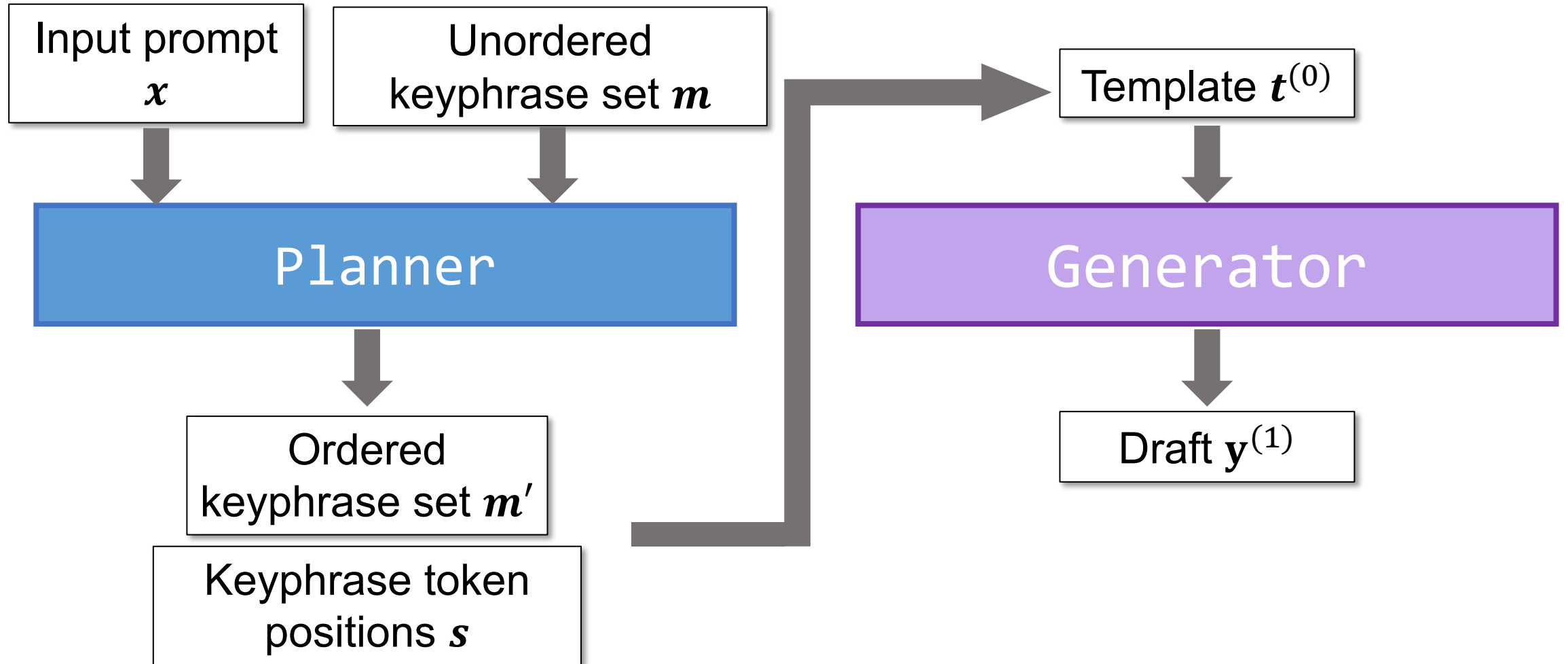


Prompt: *CMV. Donald Trump is a communist.*
Keyphrase set: {*any coherent; any topic; begin with; coherent ideology; held beliefs ...*}

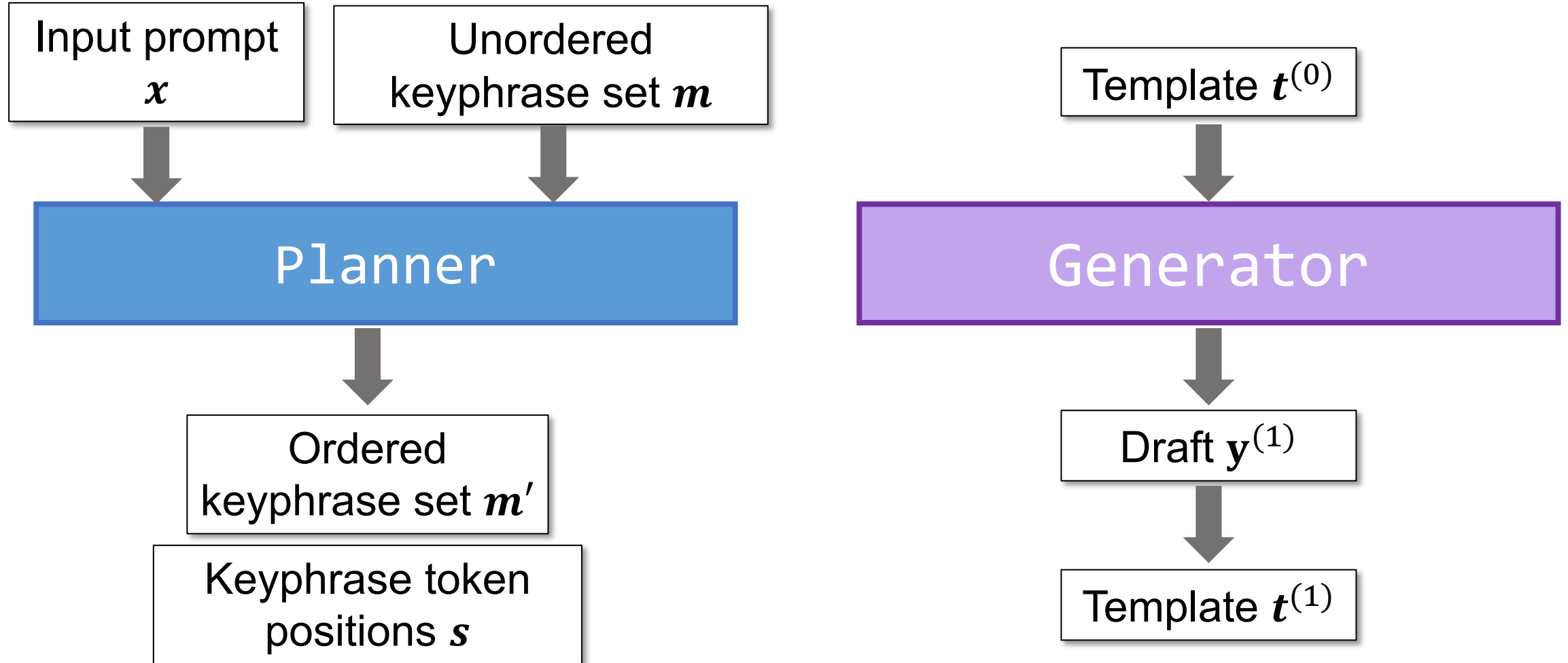
Text Plan: (a communist)₃ (begin with)₈ (coherent ideology)₁₅ ...

Template: ___ ___ ___ a communist ___
___ ___ begin with ___ ___ ...

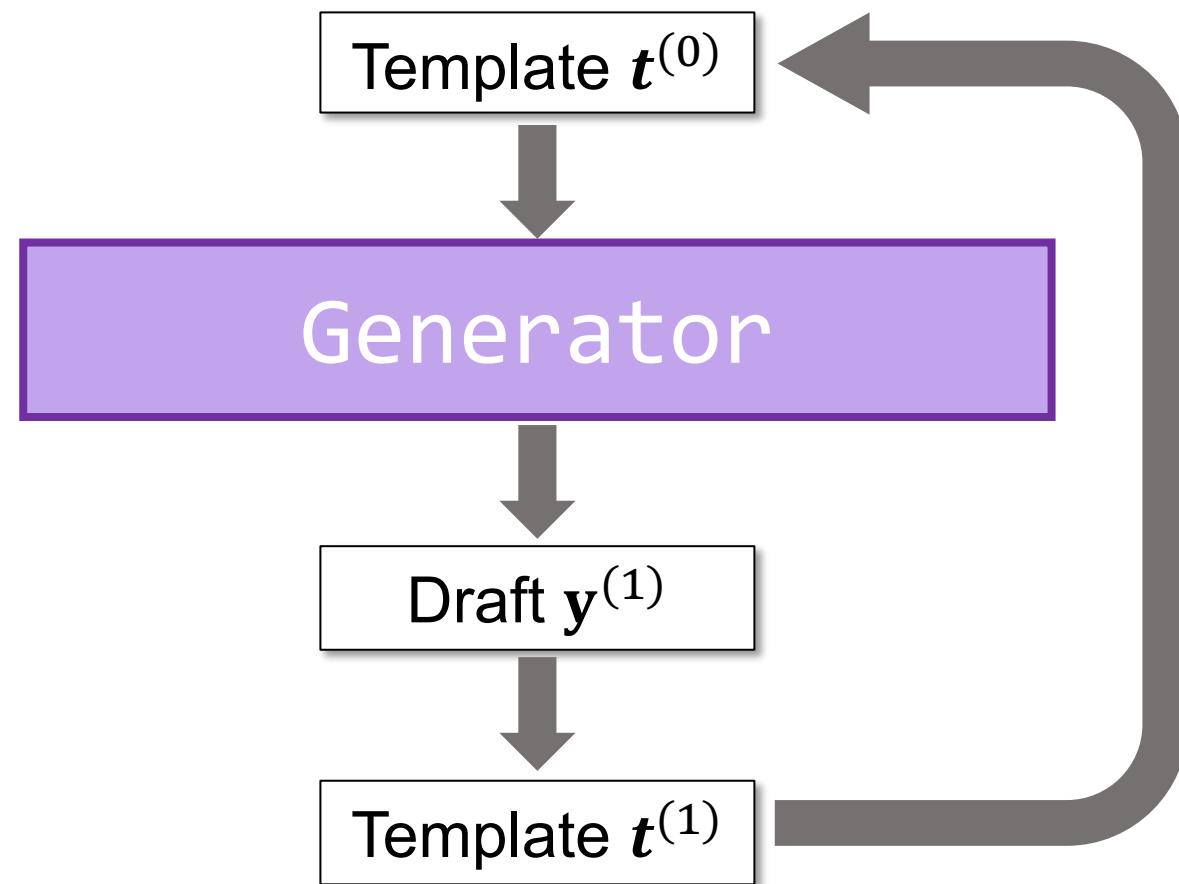
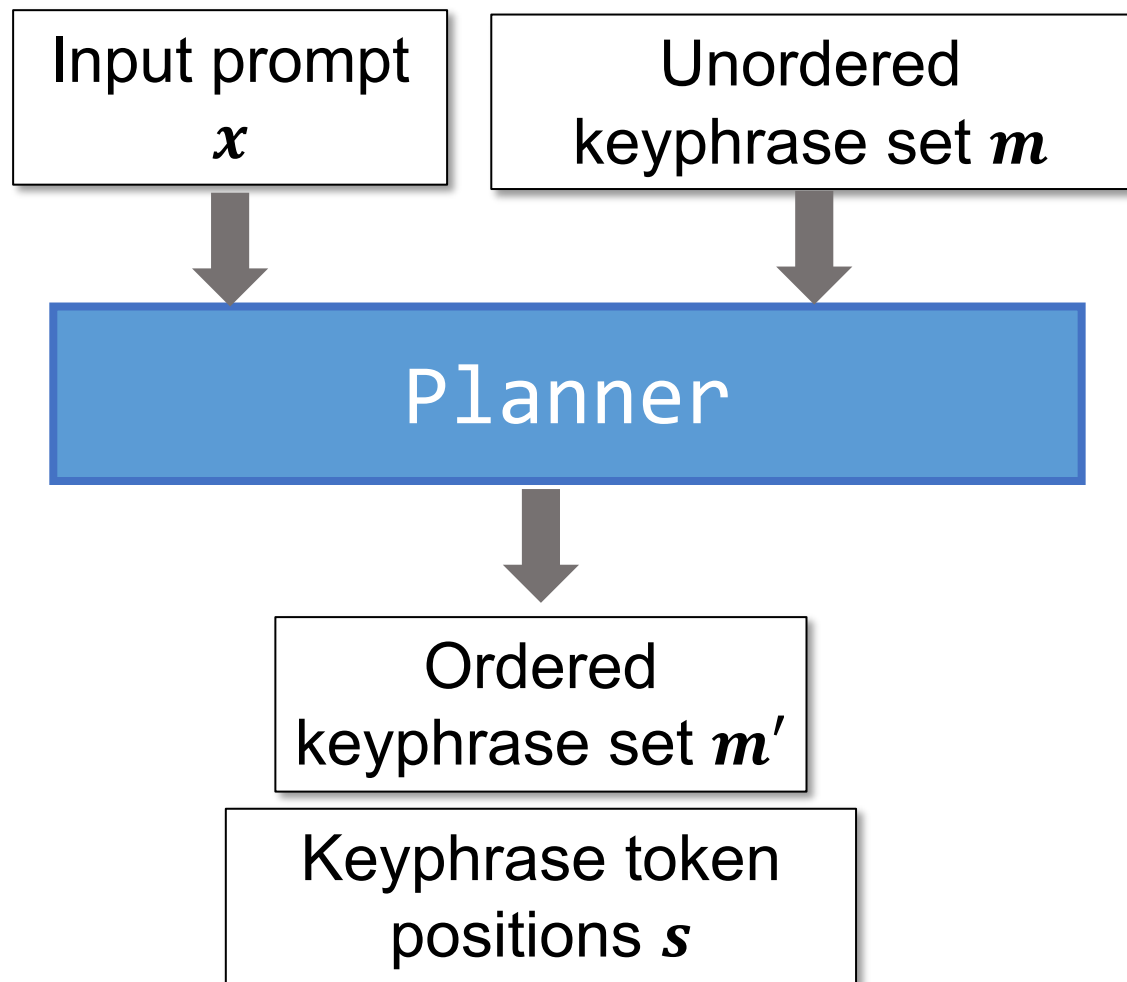
PAIR Model



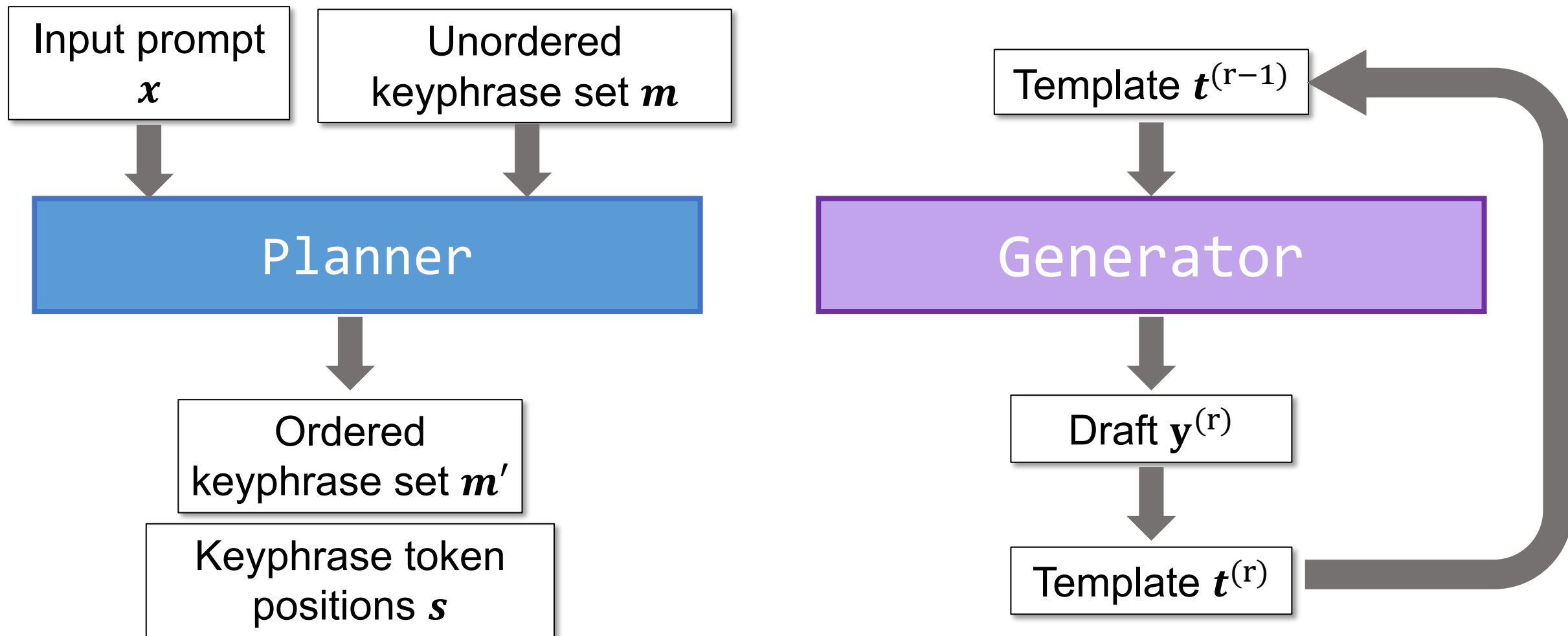
PAIR Model



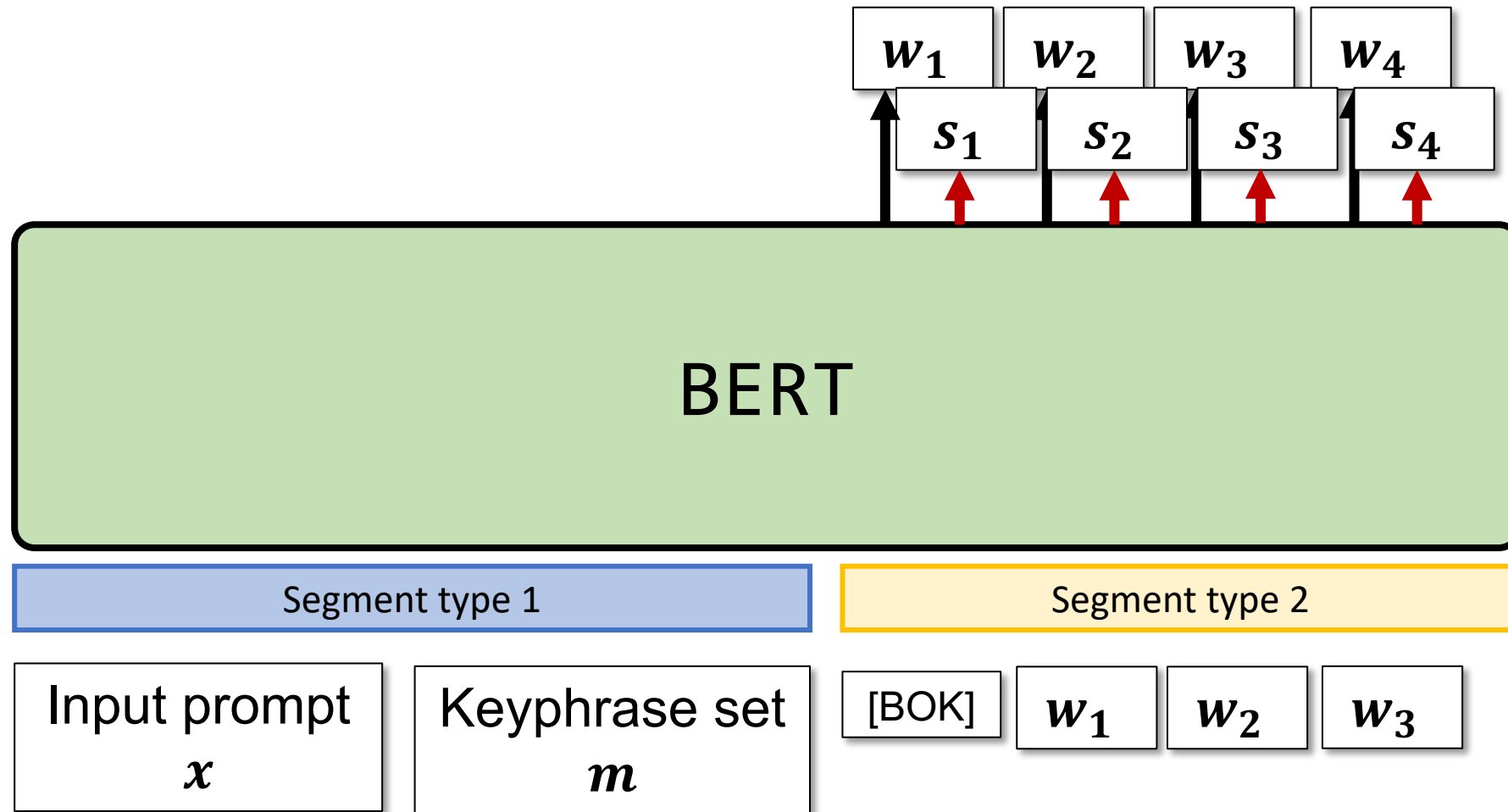
PAIR Model



PAIR Model

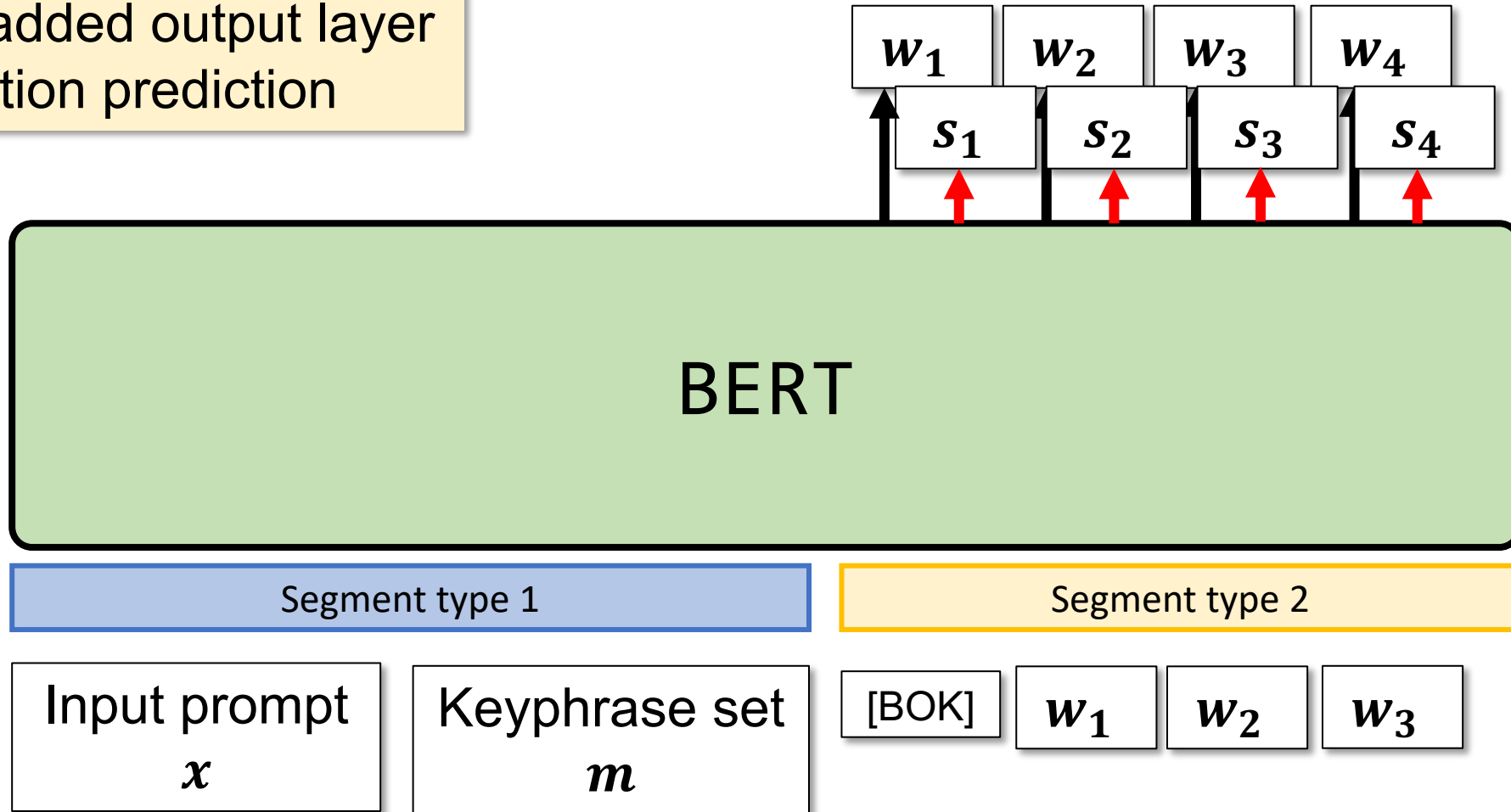


Content Planning



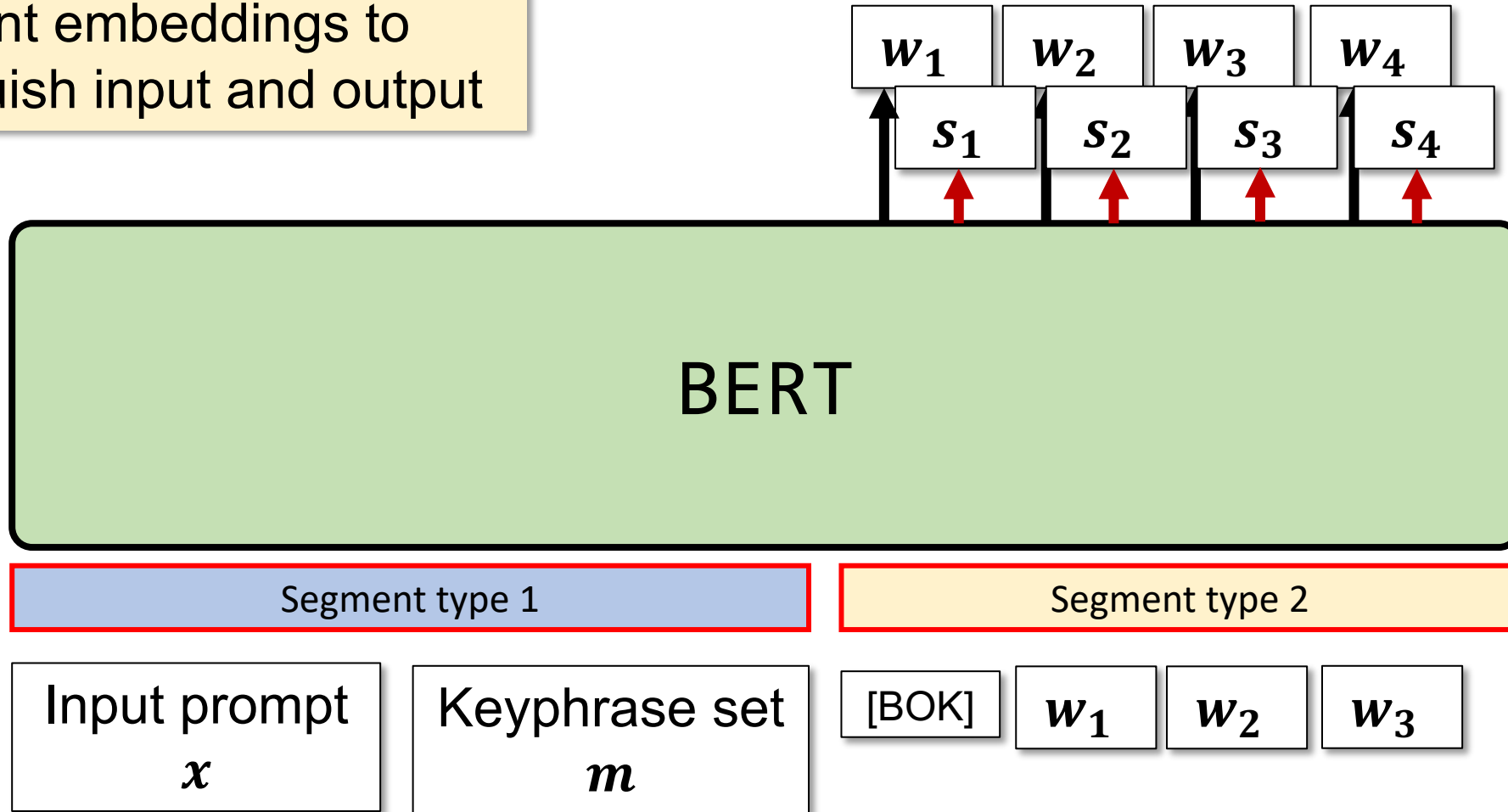
Content Planning

Newly added output layer for position prediction



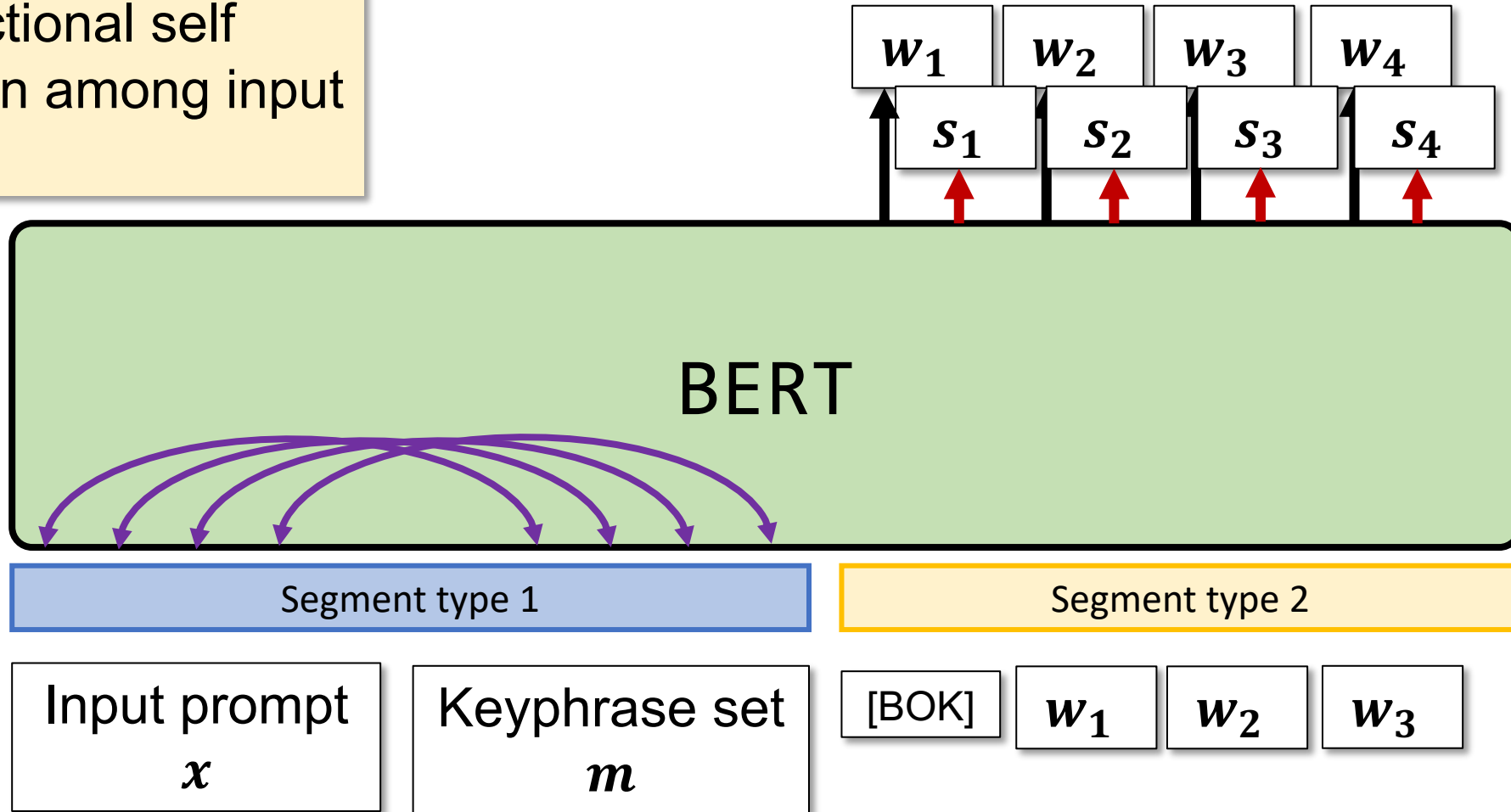
Content Planning

Segment embeddings to distinguish input and output



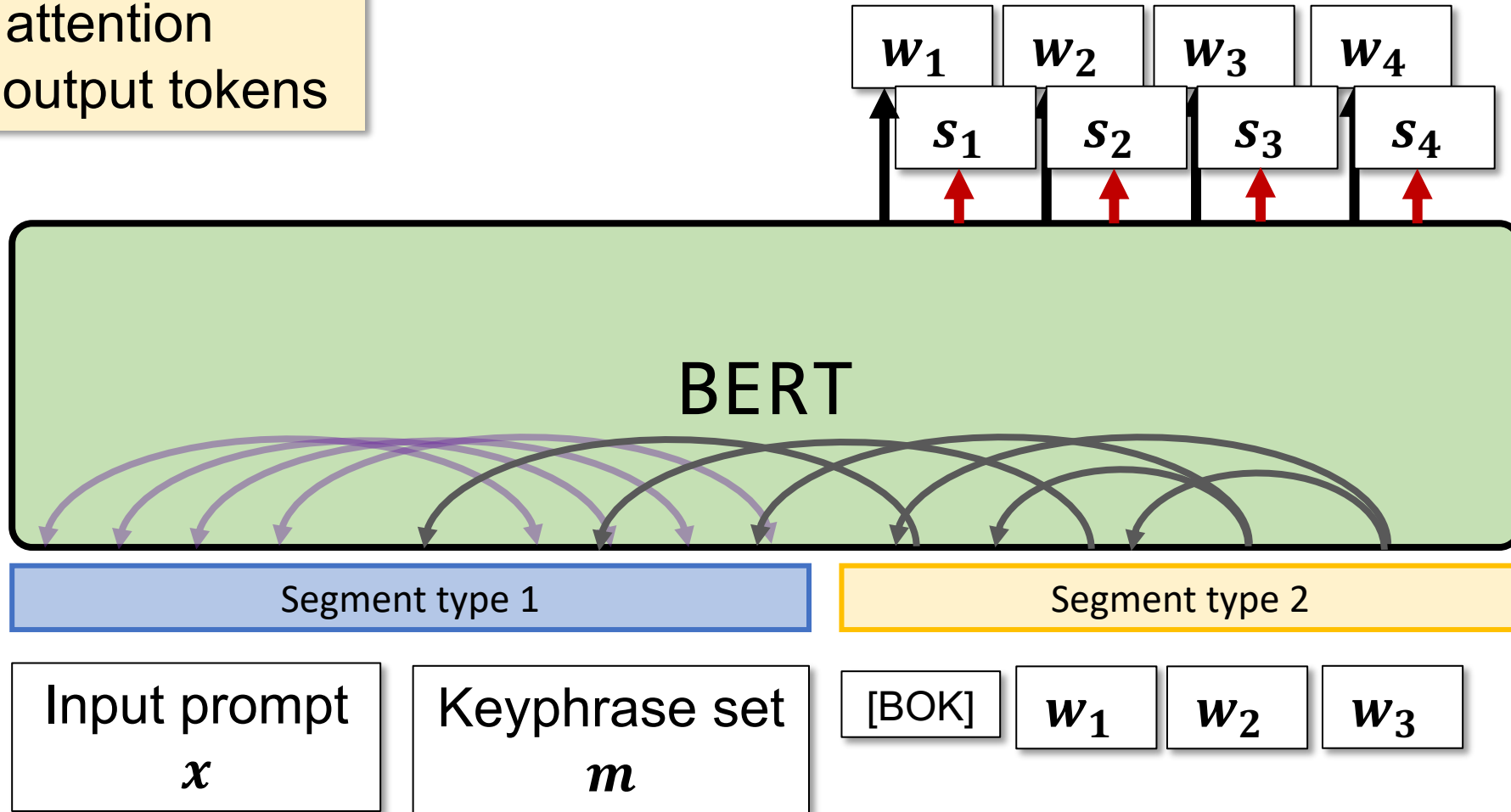
Content Planning

Bi-directional self attention among input tokens



Content Planning

Causal attention among output tokens



Mask-and-Fill

- Initial template construction

Text Plan: (a communist)₃ (begin with)₈ (coherent ideology)₁₅ ...

Mask-and-Fill

- Initial template construction

Text Plan: (a communist)₃ (begin with)₈ (coherent ideology)₁₅ ...



Template $t^{(0)}$: a communist begin with ...
 0 1 2 3 4 5 6 7 8 9 10 11

Mask-and-Fill

- Initial template construction

Text Plan: (a communist)₃ (begin with)₈ (coherent ideology)₁₅ ...

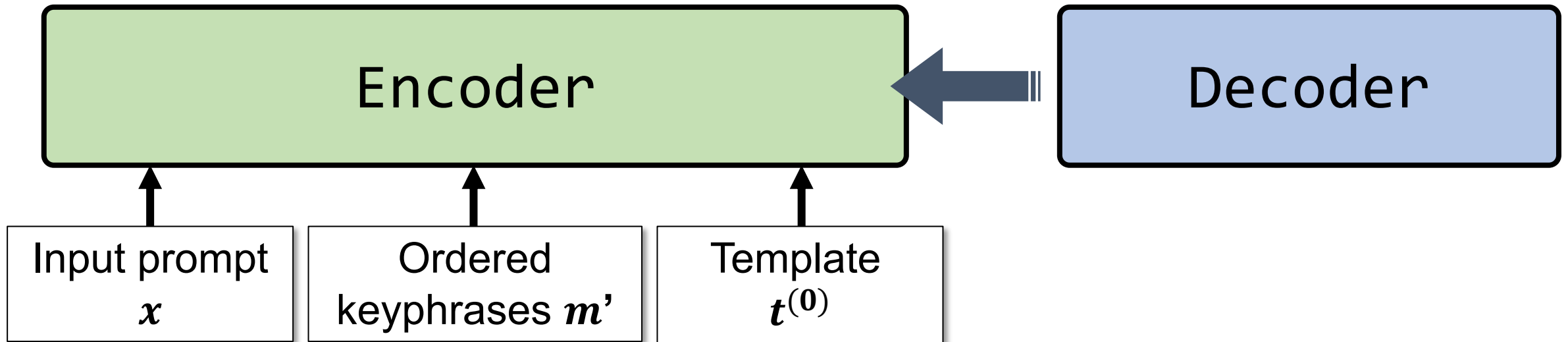


Template $t^{(0)}$: a communist begin with ...
 0 1 2 3 4 5 6 7 8 9 10 11

[MASK]

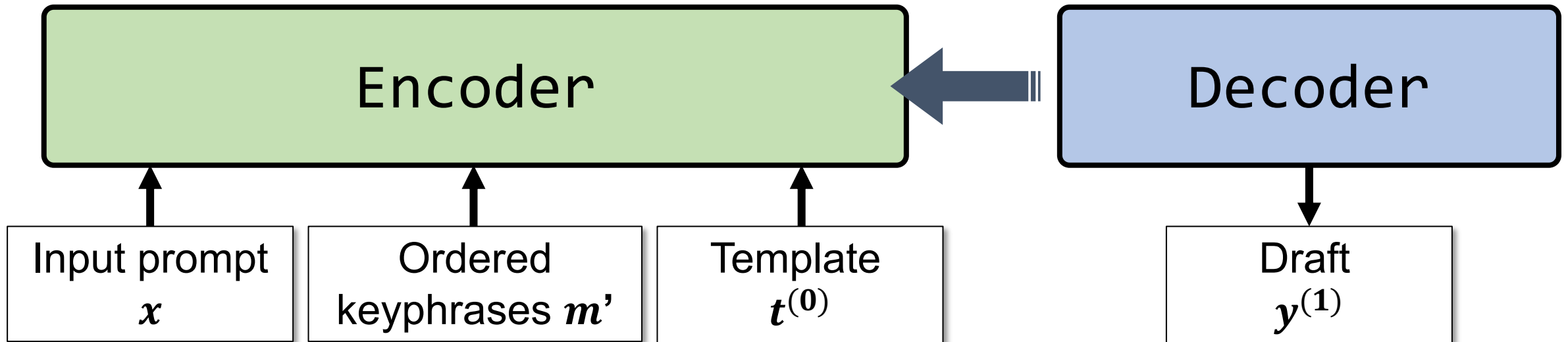
Mask-and-Fill

- BART decoding:



Mask-and-Fill

- BART decoding:



Iterative Refinement

- Single-pass output tends to have low quality, due to the noises in planning and decoding.

Iterative Refinement

- Single-pass output tends to have low quality, due to the noises in planning and decoding.
- BART is naturally a denoising model!

Iterative Refinement

- Masking tokens with low confidence [Ghazvininejad+, 2019]:

Draft $y^{(1)}$

S1: well call him a **communist** you must **begin with** that Donald Trump has **coherent ideology** to begin with.

S2: Which is unlikely.

S3: There is **no evidence** to suggest Donald Trump has **any coherent** or commonly **held beliefs** on **any topic** .

Iterative Refinement

- Masking tokens with low confidence [Ghazvininejad+, 2019]:

Draft $y^{(1)}$

S1: [mask] call him a communist you must begin with that [mask] [mask] has coherent ideology [mask] [mask] [mask].

S2: [mask] is unlikely.

S3: There is no evidence [mask] [mask] Donald Trump has any coherent or commonly held beliefs on any topic .

Iterative Refinement

- Masking tokens with low confidence [Ghazvininejad+, 2019]:

Draft $y^{(2)}$

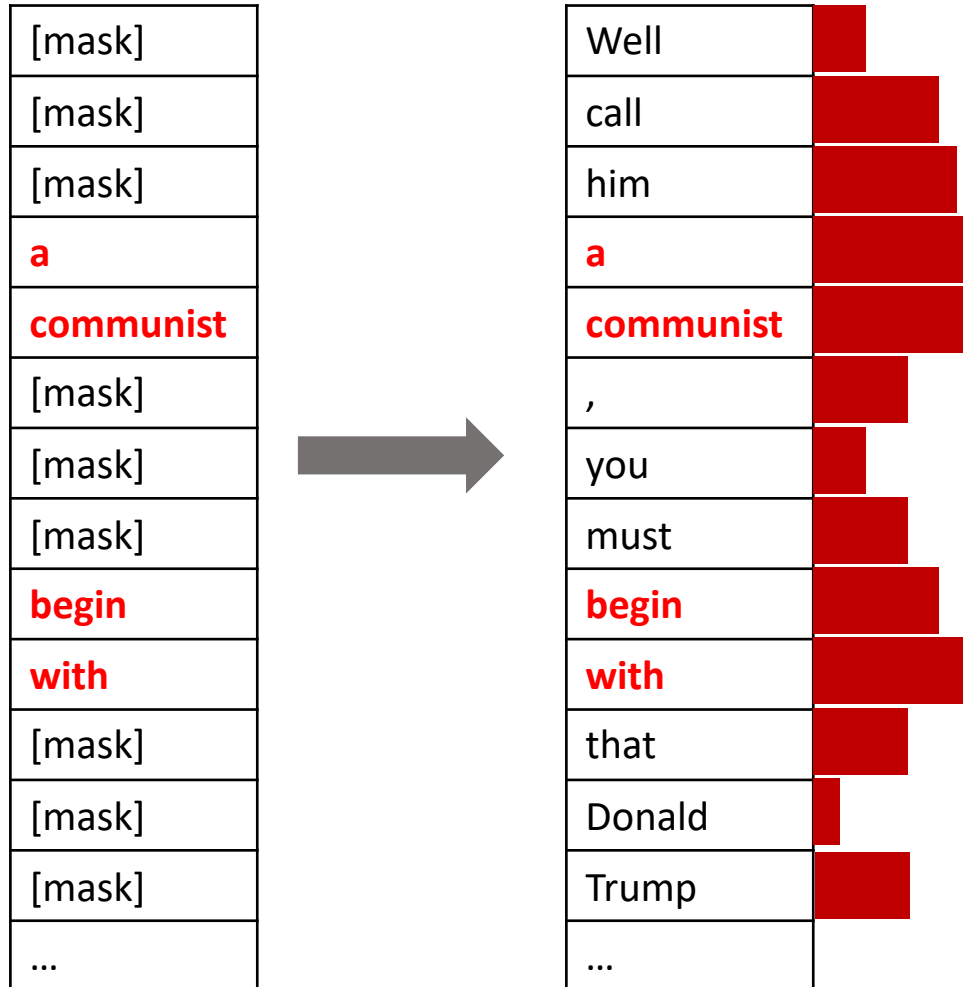
S1: *To call him a communist you must begin with that he has coherent ideology in the first place.*

S2: *That is unlikely.*

S3: *There is no evidence whatsoever that Donald Trump has any coherent or commonly held beliefs on any topic .*

Iterative Refinement

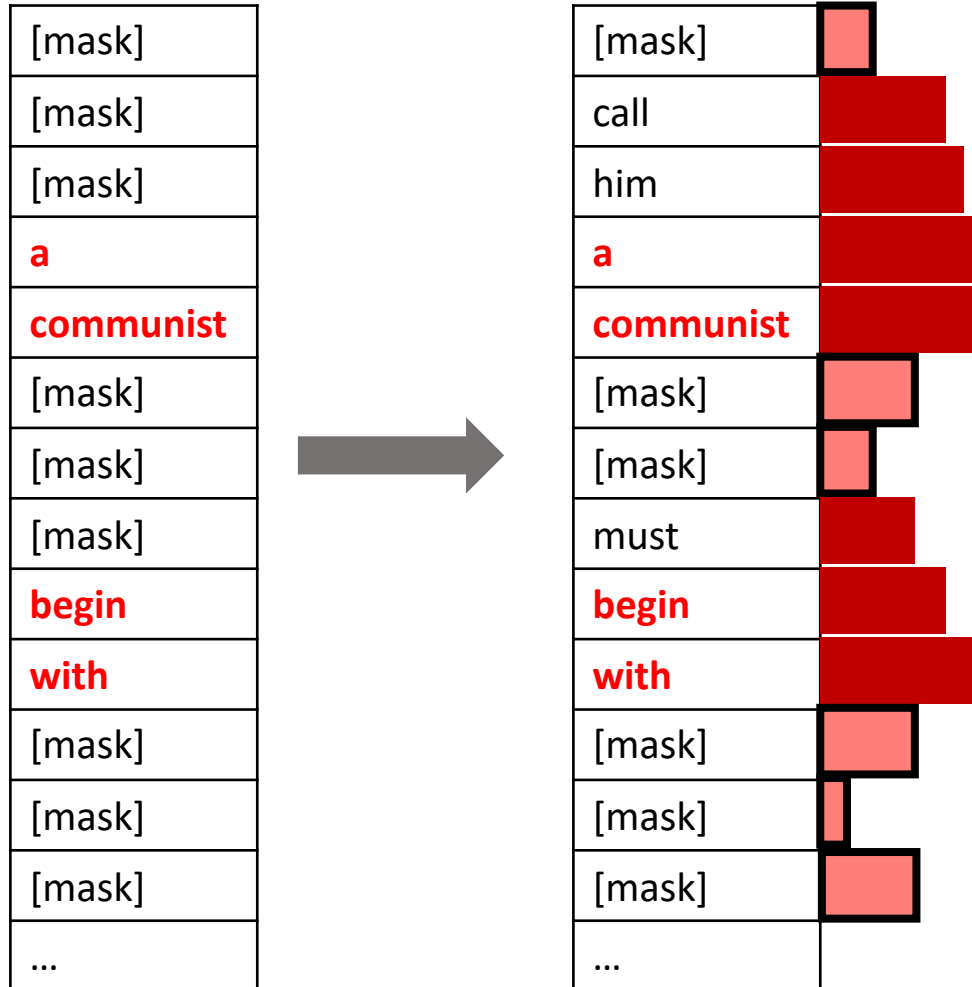
- Masking tokens with low confidence [[Ghazvininejad+, 2019](#)]:



Decoding: nucleus sampling for three times, pick the one with the lowest perplexity (by a GPT-2).

Iterative Refinement

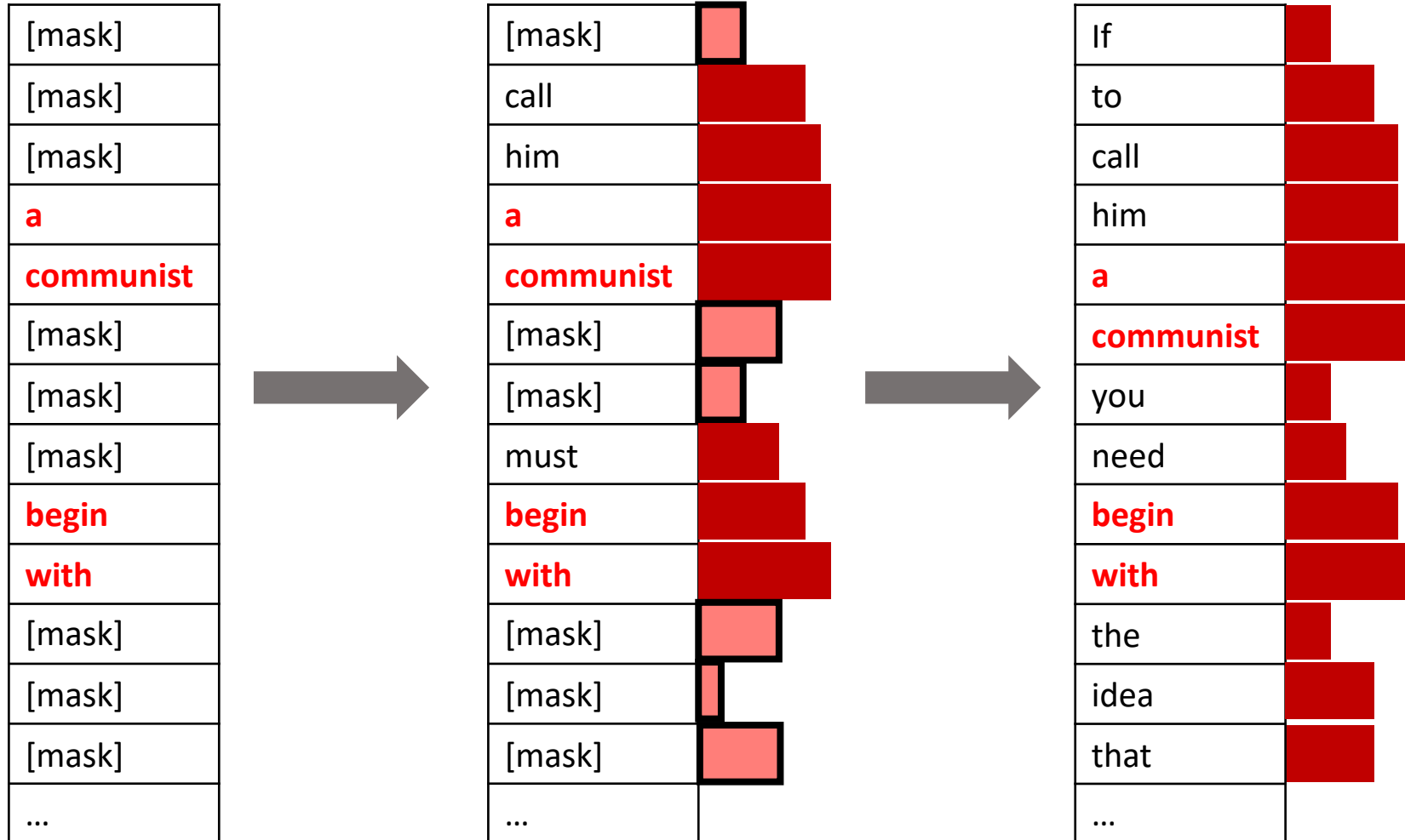
- Masking tokens with low confidence [Ghazvininejad+, 2019]:



Mask k tokens, where k linearly decrease by iterations.

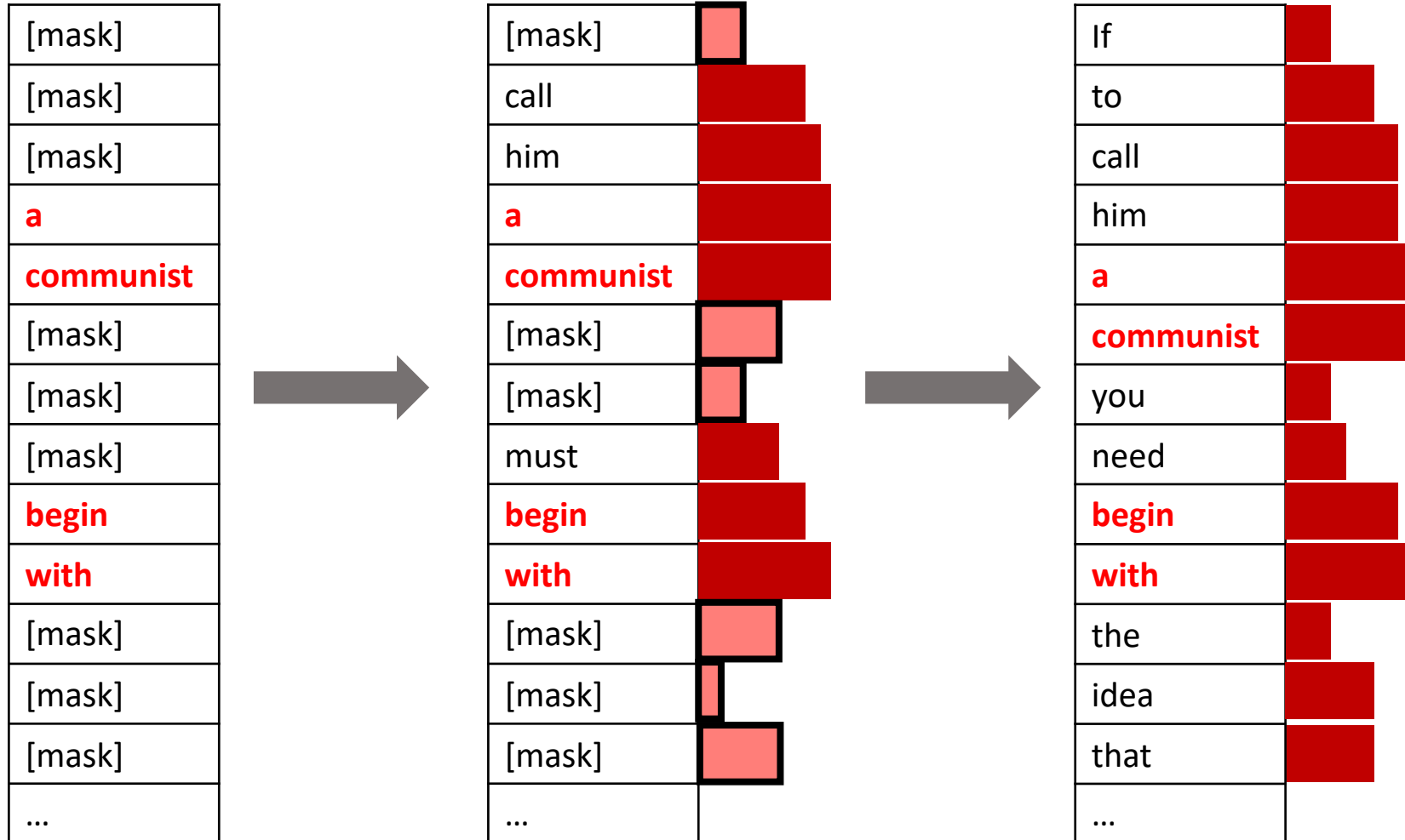
Iterative Refinement

- Masking tokens with low confidence [Ghazvininejad+, 2019]:



Iterative Refinement

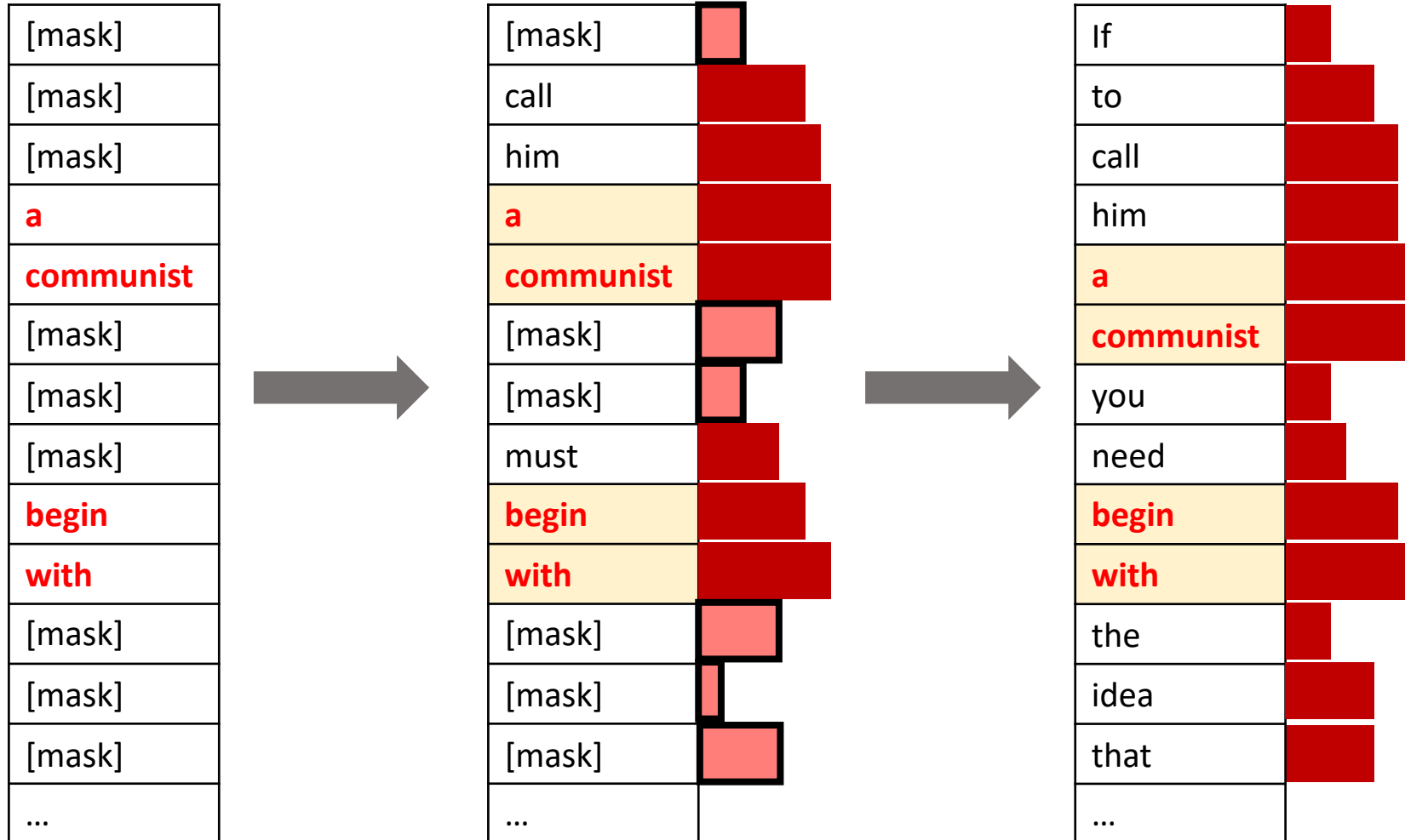
- Masking tokens with low confidence [Ghazvininejad+, 2019]:



BART is capable to *delete* and *insert* extra tokens, thus refinement is flexible.

Iterative Refinement

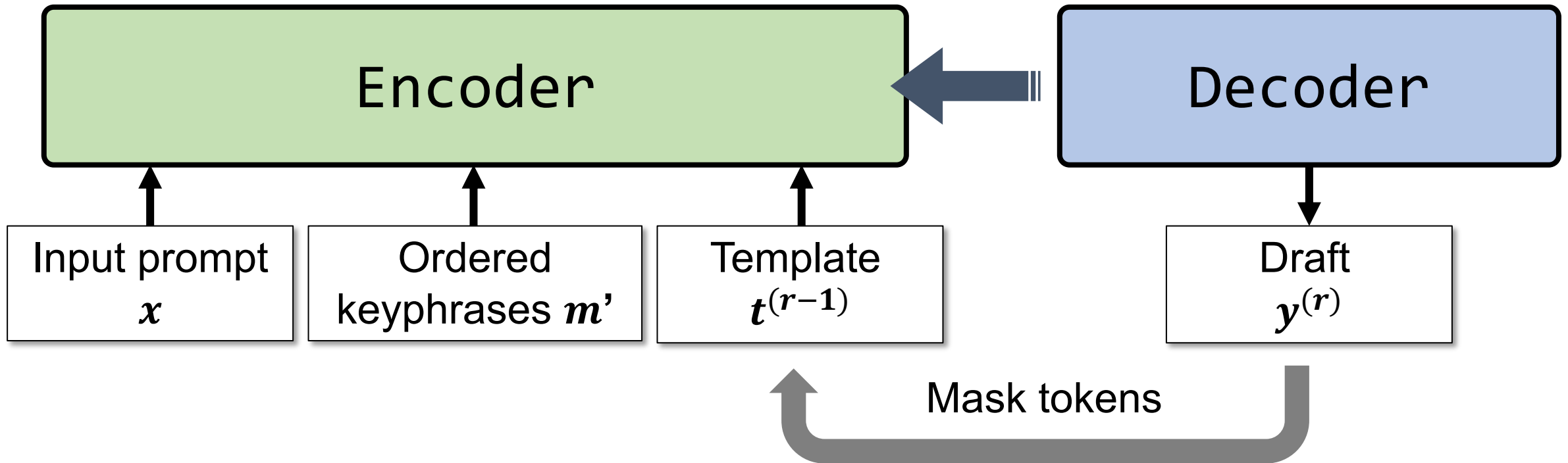
- Masking tokens with low confidence [Ghazvininejad+, 2019]:



We enforce keyphrases to be mentioned at or nearby the predicted positions.

Mask-and-Fill

- BART decoding:



Roadmap

- Motivation
- Related Work
- PAIR model
- **Experiments**
- Results
- Conclusion

Experiments

- **Tasks:**
 - Argument generation (**ARGGEN**)
 - Opinion article generation (**OPINION**)
 - News report generation (**NEWS**)

Experiments

- **Tasks:**
 - **Argument generation (ARGGEN)**

CMV: Donald Trump is a communist

Delta(s) from OP



Change My View (CMV)

r/changemyview

His international friends consists of the two most notorious communists in the modern world. Vladimir Putin member of the Communist Party of the Soviet Union (1975–91) and Xi Jinping emperor of China and General Secretary of the Communist Party of China. D.T. (Donald Trump) constantly praises China and regards Putin as a friend (Probably the terme friend is far fetched, but the fact that D.T. asked for russia to be reintegrated on the G-7 and also his one on one meetings with Putin are proof of friendship with the communist kaiser).

Experiments

- **Tasks:**

- **Argument generation (ARGGEN)**

CMV: Donald Trump is a communist

Delta(s) from OP



Change My View (CMV)

r/changemyview

His international friends consists of the two most notorious communists in the modern world. Vladimir Putin member of the Communist Party of the Soviet Union (1975–91) and Xi Jinping emperor of China and General Secretary of the Communist Party of China. D.T. (Donald Trump) constantly praises China and regards Putin as a friend (Probably the terme friend is far fetched, but the fact that D.T. asked for russia to be reintegrated on the G-7 and friendship with the communist kaiser).

↑ walking-boss 6Δ 21 points · 2 years ago

↓ Saying that he is a communist presumes to begin with that Donald Trump has some kind of coherent ideology in the first place. He doesn't. There is no evidence that he Trump has any coherent or firmly held beliefs on any topic.

↑ **White-Afro-American** 1 point · 2 years ago · *edited 2 years ago*

↓ You actually changed my view to Donald Trump acts like a communist.

Thank you for your answer. Δ

Experiments

- Tasks:

- Argument generation (**ARGGEN**)

CMV: Donald Trump is a communist

Delta(s) from OP

OP -> input prompt



Change My View (CMV)

r/changemyview

His international friends consists of the two most notorious communists in the modern world. Vladimir Putin member of the Communist Party of the Soviet Union (1975-91) and Xi Jinping emperor of China and General Secretary of the Communist Party of China. D.T. (Donald Trump) constantly praises China and regards Putin as a friend (Probably the terme friend is far fetched, but the fact that D.T. asked for russia to be reintegrated on the G-7 and friendship with the communist kaiser).

↑ walking-boss 6Δ 21 points · 2 years ago

↓ Saying that he is a communist presumes to begin with that Donald Trump has some kind of coherent ideology in the first place. He doesn't. There is no evidence that he Trump has any coherent or firmly held beliefs on any topic.

↑ White-Afro-American 1 point · 2 years ago · edited 2 years ago

↓ You actually changed my view to Donald Trump acts like a communist.

Thank you for your answer. Δ

High quality replies -> Keyphrases, target

Experiments

- Tasks:
 - Argument generation (**ARGGEN**)

CMV: Donald Trump is a communist

Delta(s) from OP

His international
Putin member
and General Se
and regards Pu
russia to be rei
friendship with the communist kaiser).

Size	Prompt Len.	# Keyphrases	Target Len.
56,504	19.4	20.6	116.6

Saying that he is a communist presumes to begin with that Donald Trump has some kind of coherent ideology in the first place. He doesn't. There is no evidence that he Trump has any coherent or firmly held beliefs on any topic.

↑ **White-Afro-American** 1 point · 2 years ago · edited 2 years ago

↓ You actually changed my view to Donald Trump acts like a communist.

Thank you for your answer. Δ

Experiments

- **Tasks:**

- Argument generation (**ARGGEN**)
- Opinion article generation (**OPINION**)
- News report generation (**NEWS**)

Articles from the New York Times corpus [[Sandhaus, 2008](#)].

“Top/Opinion” -> **OPINION**
“Top/News” -> **NEWS**



Experiments

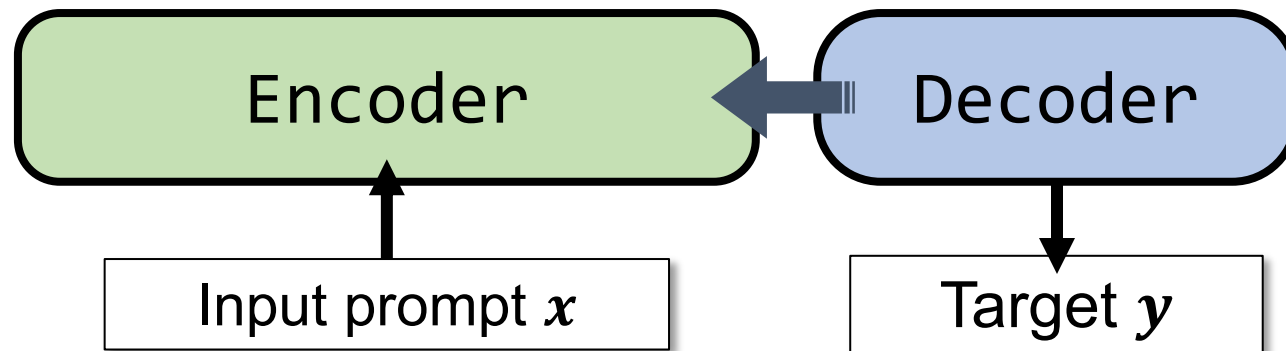
- **Tasks:**
 - Argument generation (**ARGGEN**)
 - Opinion article generation (**OPINION**)
 - News report generation (**NEWS**)

	Size	Prompt Len.	# Keyphrases	Target Len.
OPINION	104,610	6.1	19.0	205.6
NEWS	239,959	7.0	30.3	282.7

Experiments

- **Baselines:**

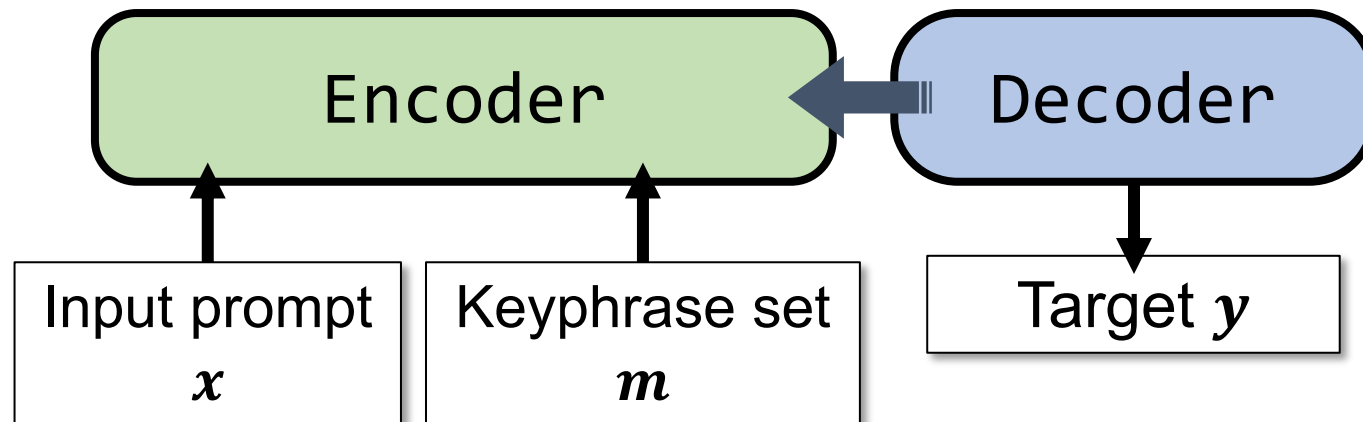
- SEQ2SEQ: directly generates target from the prompt
- KPSEQ2SEQ: generates target from prompt and unordered keyphrase set



Experiments

- **Baselines:**

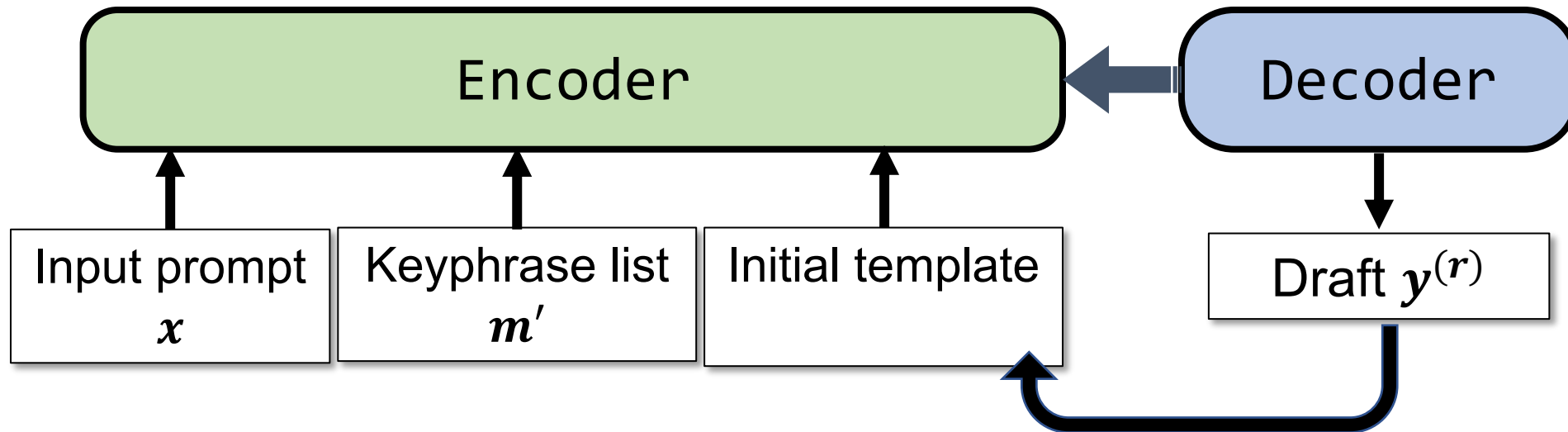
- SEQ2SEQ: directly generates target from the prompt
- KPSEQ2SEQ: generates target from prompt and *unordered* keyphrase set



Experiments

- Comparisons:

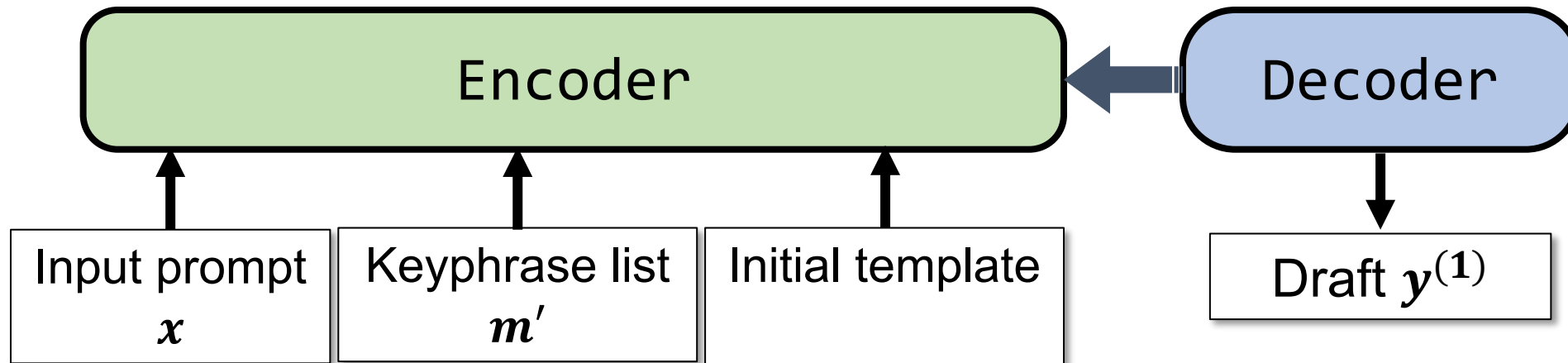
- PAIR_{full} : our proposed model with iterative refinement



Experiments

- Comparisons:

- PAIR_{full} : our proposed model with iterative refinement
- PAIR_{full} without refinement: one-pass generation

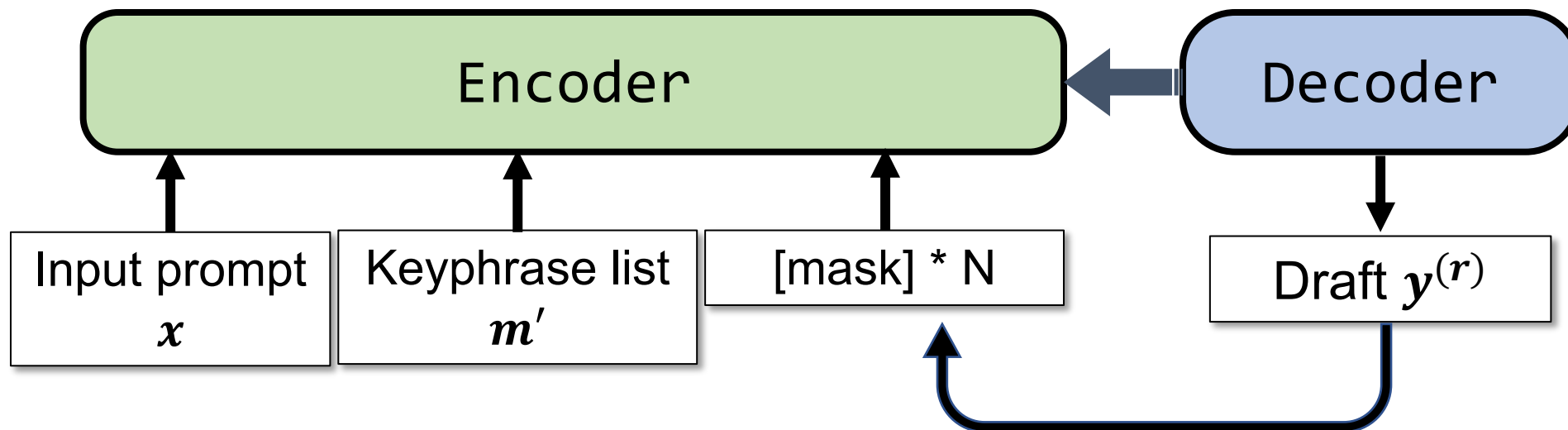


Experiments

- Comparisons:

- PAIR_{full}: our proposed model with iterative refinement
- PAIR_{full} without refinement: one-pass generation
- PAIR_{light}: replace initial template with all masks

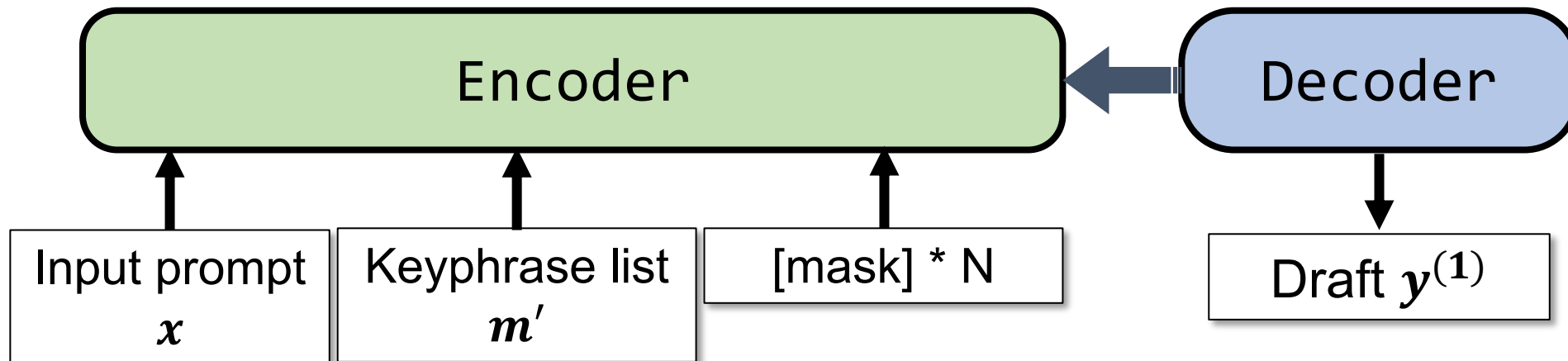
Model without keyphrase position information.



Experiments

- Comparisons:

- PAIR_{full} : our proposed model with iterative refinement
- PAIR_{full} without refinement: one-pass generation
- PAIR_{light} : replace initial template with all masks
- PAIR_{light} without refinement



Roadmap

- Motivation
- Related Work
- PAIR model
- Experiments
- **Results**
- Conclusion

Results

- Automatic Evaluation – **ARGGEN**

Results with ground-truth content plans.

BLEU-4

ROUGE-L

METEOR

Length

Results

- Automatic Evaluation – **ARGGEN**

	BLEU-4	ROUGE-L	METEOR	Length
SEQ2SEQ	0.76	13.80	9.36	97
KPSEQ2SEQ	6.78	19.43	15.98	97

Results

• Automatic Evaluation – **ARGGEN**

	BLEU-4	ROUGE-L	METEOR	Length
SEQ2SEQ	0.76	13.80	9.36	97
KPSEQ2SEQ	6.78	19.43	15.98	97
PAIR _{light}	26.38	47.97	31.64	119
PAIR _{light} w/o refine	25.17	46.84	31.31	120
PAIR _{full}	36.09	56.86	33.30	102
PAIR _{full} w/o refine	34.09	55.42	32.74	101

Results

- Automatic Evaluation – **ARGGEN**

Position information helps (full model).

	BLEU-4	ROUGE-L	METEOR	Length
SEQ2SEQ	0.76	13.80	9.36	97
KPSEQ2SEQ	6.78	19.43	15.98	97
PAIR _{light}	26.38	47.97	31.64	119
PAIR _{light} w/o refine	25.17	46.84	31.31	120
PAIR _{full}	36.09	56.86	33.30	102
PAIR _{full} w/o refine	34.09	55.42	32.74	101

Results

• Automatic Evaluation – **ARGGEN**

Refinement improves all metrics.

	BLEU-4	ROUGE-L	METEOR	Length
SEQ2SEQ	0.76	13.80	9.36	97
KPSEQ2SEQ	6.78	19.43	15.98	97
PAIR _{light}	26.38 ↑	47.97 ↑	31.64 ↑	119
PAIR _{light} w/o refine	25.17	46.84	31.31	120
PAIR _{full}	36.09 ↑	56.86 ↑	33.30 ↑	102
PAIR _{full} w/o refine	34.09	55.42	32.74	101

Results

- Automatic Evaluation – **OPINION**

	BLEU-4	ROUGE-L	METEOR	Length
SEQ2SEQ	1.42	15.97	10.97	156
KPSEQ2SEQ	11.38	22.75	18.38	164
PAIR _{light}	16.27 ↑	33.30 ↑	24.32 ↑	210
PAIR _{light} w/o refine	15.45	32.35	24.11	214
PAIR _{full}	23.12 ↑	40.53 ↑	24.73 ↑	167
PAIR _{full} w/o refine	22.17	39.71	24.65	169

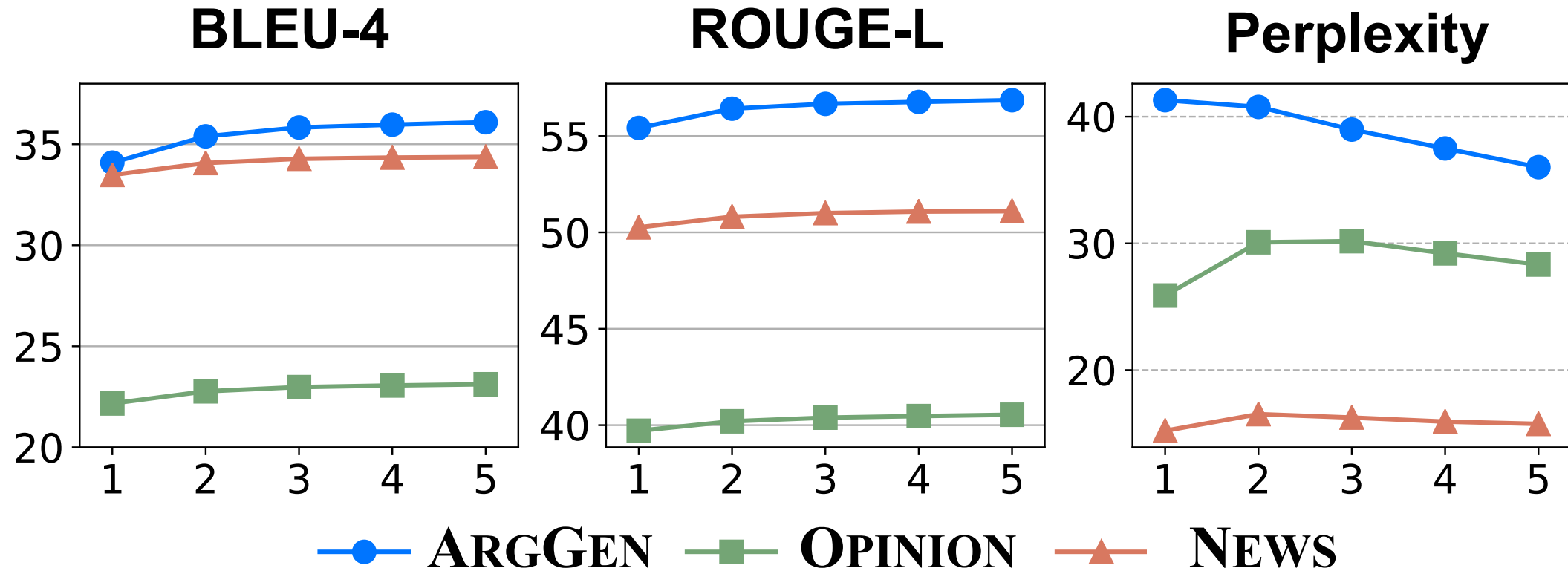
Results

- Automatic Evaluation – NEWS

	BLEU-4	ROUGE-L	METEOR	Length
SEQ2SEQ	1.11	15.60	10.10	242
KPSEQ2SEQ	11.61	21.05	18.61	286
PAIR _{light}	28.03 ↑	43.39 ↑	27.70 ↑	272
PAIR _{light} w/o refine	27.32	43.08	27.35	278
PAIR _{full}	34.37 ↑	51.10 ↑	29.50 ↑	259
PAIR _{full} w/o refine	33.48	50.27	29.26	260

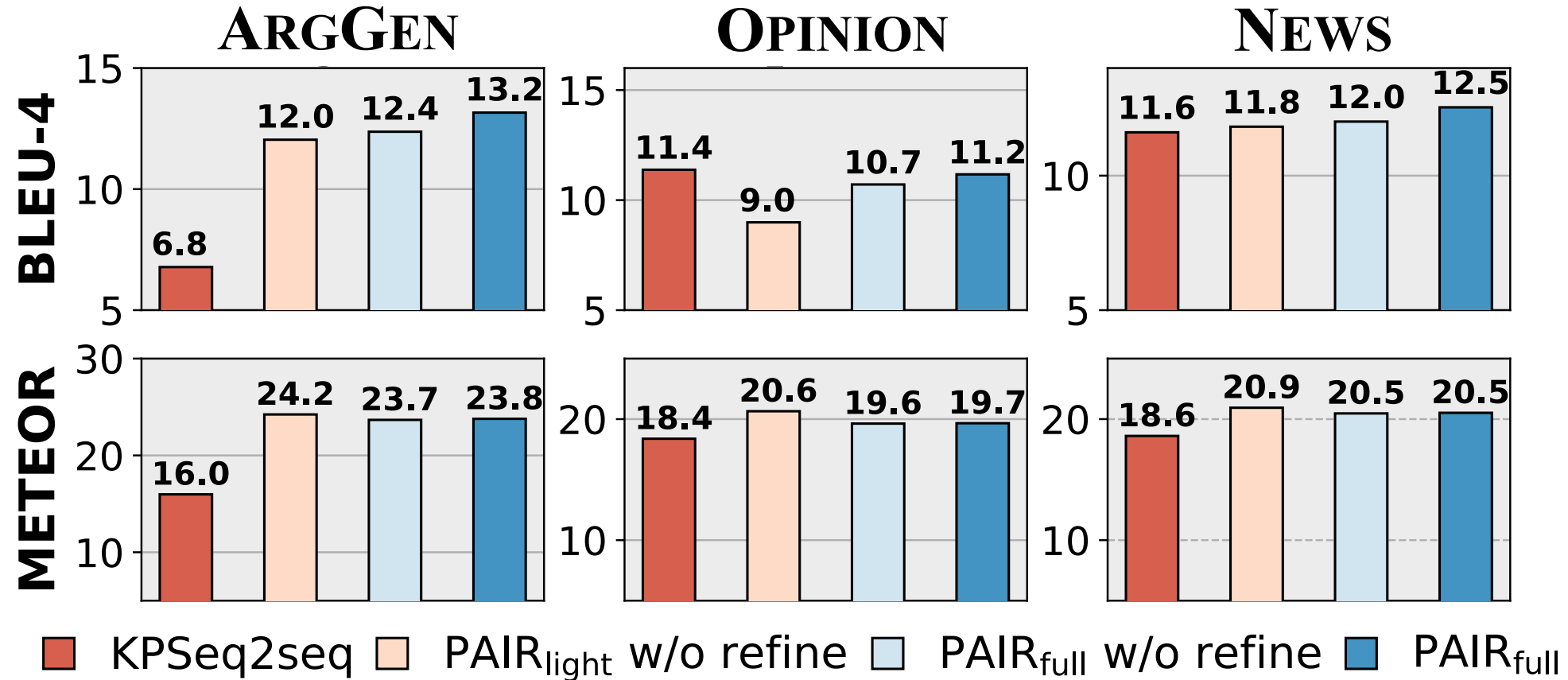
Results

- Change across iterations:



Results

- Results with system predicted content plans:



Results

- **Human Evaluation**

- **Fluency** (1-5): if the output is grammatical and readable
- **Coherence** (1-5): if the information organization is natural and logical
- **Relevance** (1-5): if the topic is related to the prompt and whether the stance is correct

Results

- **Human Evaluation**

- **Fluency** (1-5): if the output is grammatical and readable
- **Coherence** (1-5): if the information organization is natural and logical
- **Relevance** (1-5): if the topic is related to the prompt and whether the stance is correct

	Fluency	Coherence	Relevance
KPSEQ2SEQ	4.63	3.28	2.79
PAIR _{light}	4.75	3.97	3.85
PAIR _{full}	4.46	3.76	3.79

Sample Output - ArgGen

Prompt: CMV. The EU needs its own independent armed forces

I don't see **the primary benefit** of **a military** under **the authority** of **EU institutions**. Instead, I think **the EU** should set up its own **armed forces** with **UN Security Council authorisation**, to **the EU member states**, or **any subset** of EU member states that wishes not to participate. Ultimately, **the primary advantage** of **a EU military** is that **military operations conducted under the authority** of EU institutions already have **a level of legitimacy** comparable to **individual national states**, [...]

Sample Output - Opinion

Prompt: The World's Best Job

I was happy to see that **travel writers** are getting better. Many of **those travel writers** and **research guides** in **Goa** and **Romania** had a **rough time** there, but the vast majority of them, as is **all travel writers**. They all work hard and do a lot of **research trips** all over the world. Their research trips, are **hectic**, and they are much more than **the ideal job** or **hobby**. [...]

Sample Output - News

Prompt: McGarry Bowen Added To Kraft Agency Roster

Kraft Foods expanded its **agency roster** by **naming McGarry Bowen** in New York to create campaigns for **salad dressings, mayonnaise** and **barbecue sauces bearing the Kraft brand name**. Spending was estimated at \$10 million. The decision to expand its relationship with McGarry Bowen is part of a trend in which **big marketers** like big marketers like **Kraft shift creative assignments** from small, **independent agencies** to **larger agencies** at the same time. [...]

What is updated during refinement?

- About 85% refinement operations changed content, 1~5% added tokens, 2~6% deleted tokens.

What is updated during refinement?

- About 85% refinement operations changed content, 1~5% added tokens, 2~6% deleted tokens.
- **38%~41% masks are put on unigrams, 22% on bigrams, 13~15% on trigrams.**

What is updated during refinement?

- Top frequent changes:

ARGGEN		OPINION		NEWS	
Old	New	Old	New	Old	New
because	,	Feb	Jan	added	said
It	That	Feb	Dec	with	and
?	.	is	was	from	to
their	your	does	is	did	would
need	have	from	of	announced	said
actually	really	more	less	While	Although
may	would	can	cannot	first	second

What is updated during refinement?

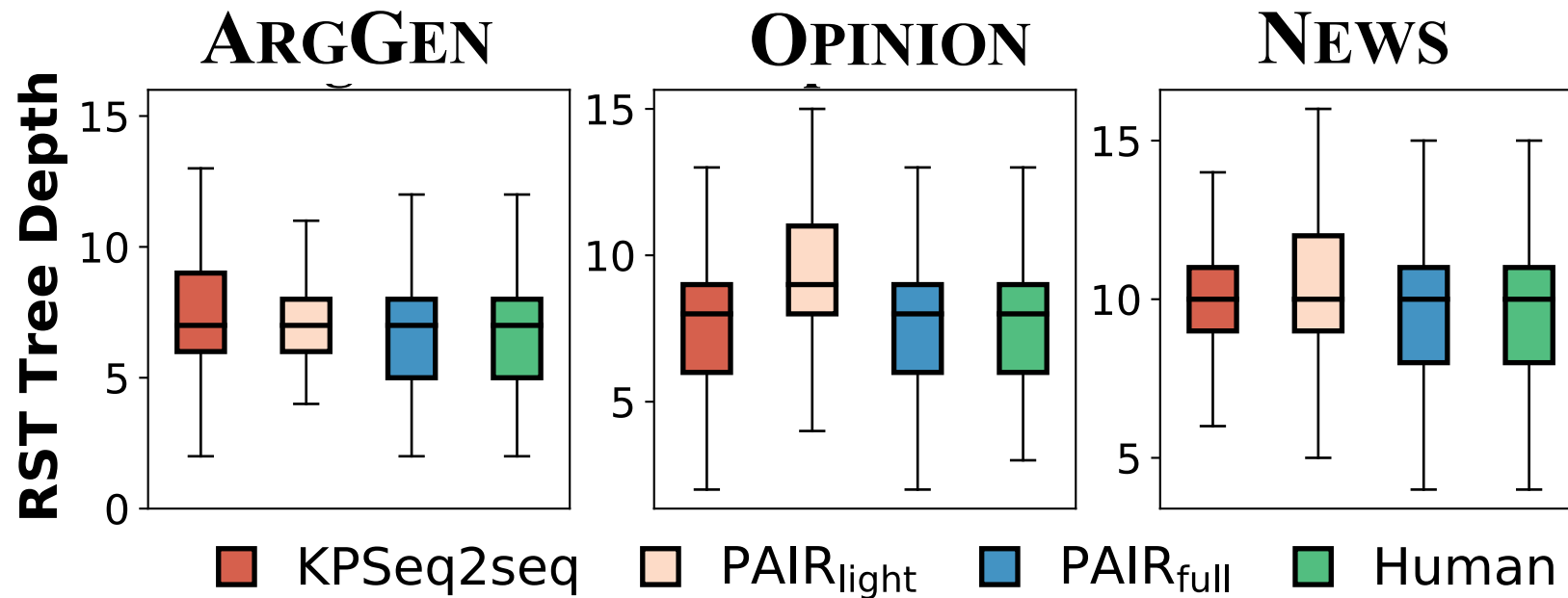
Refinement can produce factual errors!

- Top frequent changes:

ARGGEN		OPINION		NEWS	
Old	New	Old	New	Old	New
because	,	Feb	Jan	added	said
It	That	Feb	Dec	with	and
?	.	is	was	from	to
their	your	does	is	did	would
need	have	from	of	announced	said
actually	really	more	less	While	Although
may	would	can	cannot	first	second

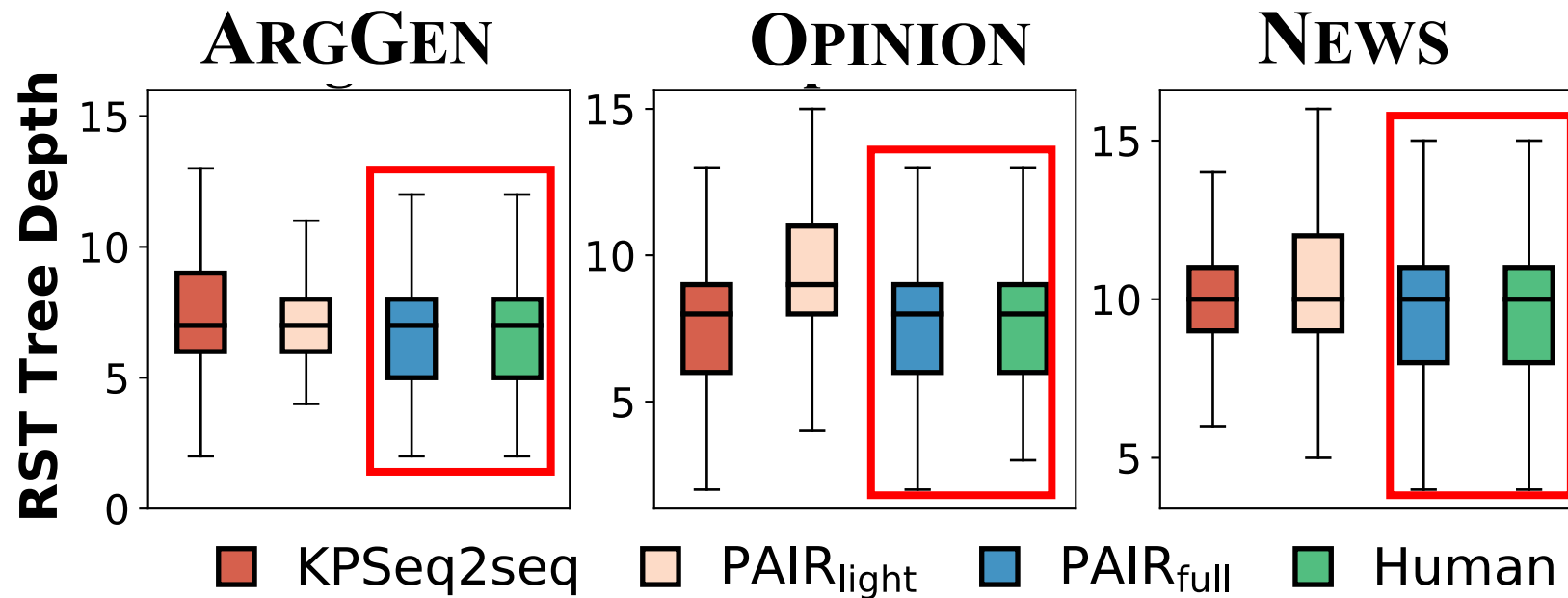
Is PAIR better at capturing discourse?

- Study on RST tree depth using DPLP [Ji & Eisenstein, 2014]



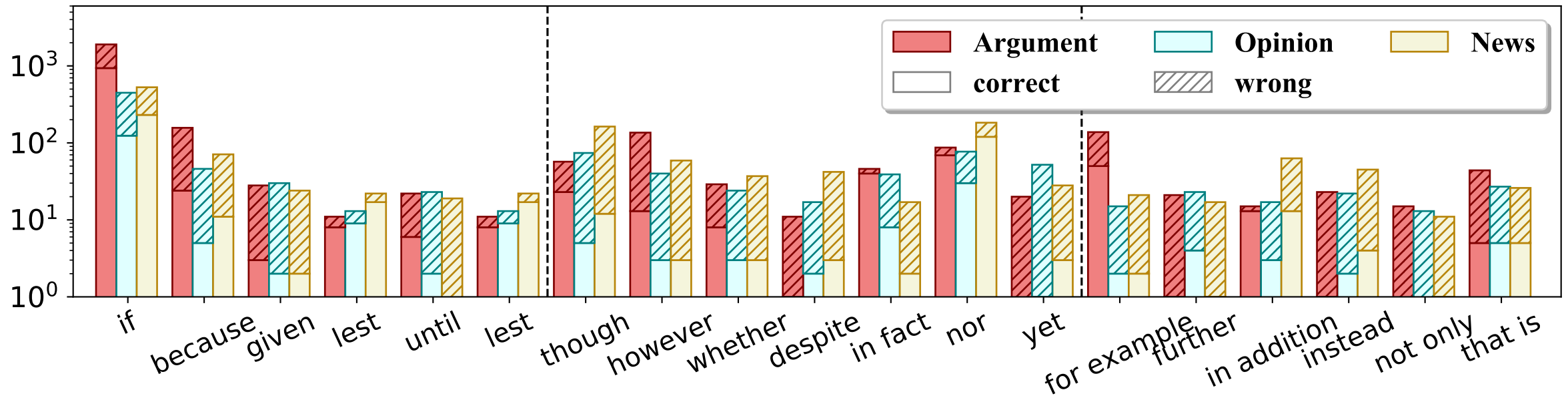
Is PAIR better at capturing discourse?

- Study on RST tree depth using DPLP [Ji & Eisenstein, 2014]



Is PAIR better at capturing discourse?

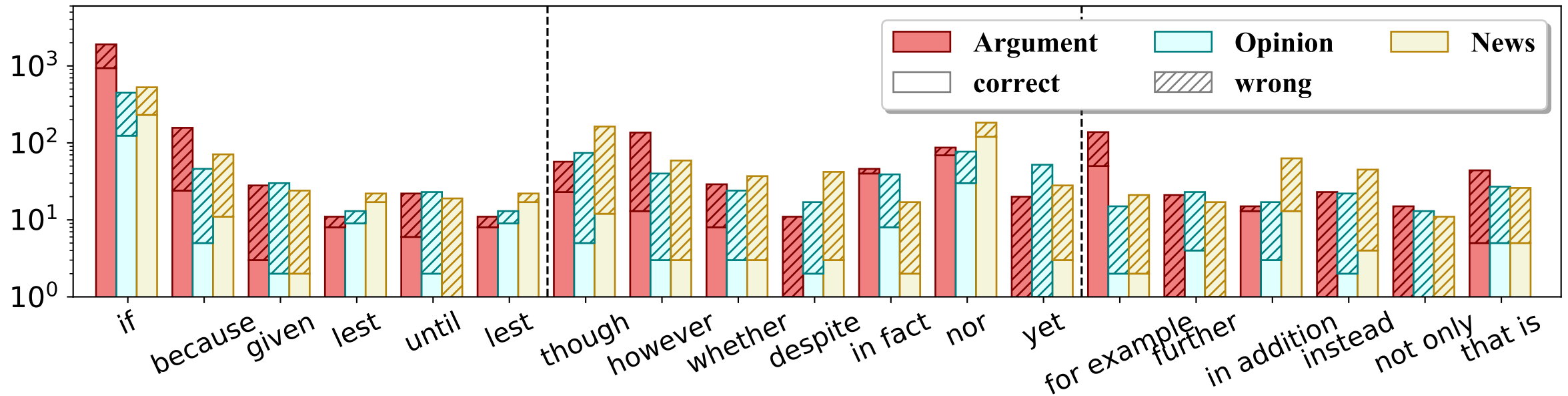
- Discourse marker accuracy



Is PAIR better at capturing discourse?

- Discourse marker accuracy

PAIR is better at Contingency than Comparison and Expansion.



Conclusion

- We propose a content-controlled generation framework with separate planning and realization.
- Our model leverages the masked language model for iterative refinement to improve the output quality.
- Both automatic and human evaluation show the effectiveness of our model over three tasks.

Thanks!



<https://xinyuhua.github.io/Resources/emnlp20/>



hua.x@northeastern.edu



<https://github.com/XinyuHua/pair-emnlp2020>